



A Poisson random walk model of response times

Blurton, Steven; Kyllingsbæk, Søren; Nielsen, Carsten Søren; Bundesen, Claus

Published in:
Psychological Review

Publication date:
2020

Document version
Peer reviewed version

Document license:
[Unspecified](#)

Citation for published version (APA):
Blurton, S., Kyllingsbæk, S., Nielsen, C. S., & Bundesen, C. (2020). A Poisson random walk model of response times. *Psychological Review*, 127(3), 362-411. <http://doi.org/10.1037/rev0000179>

A Poisson random walk model of response times

Steven P. Blurton^a, Søren Kyllingsbæk^a, Carsten S. Nielsen^{a,b}, & Claus Bundesen^a

^a Department of Psychology, University of Copenhagen, Denmark

^b Department of Economics, University of Copenhagen, Denmark

Correspondence

Steven P. Blurton

Department of Psychology

University of Copenhagen

Øster Farimagsgade 2A

1353 København K

Denmark

E-mail: steven.blurton@psy.ku.dk

Abstract: 236 words

Manuscript: 39603 words, 16 figures, 10 tables

The formulas for the random walk model of response times were adapted from Feller (1968) by Bundesen during Bundesen's stay as visiting scholar at Stanford University in 1977 (see Bundesen, 1982). Parts of the data and different aspects of the theory described in this paper have been presented at various conferences (European Mathematical Psychology Group meeting 2015, 2016; Meeting of the Vision Science Society, 2016; Meeting of the Society of Mathematical Psychology, 2017; European Conference on Visual Perception, 2017).

Abstract

Based on the simple *what first comes to mind* rule, the theory of visual attention (TVA; Bundesen, 1990) provides a comprehensive account of visual attention that has been successful in explaining performance in visual categorization for a variety of attention tasks. If the stimuli to be categorized are mutually confusable, a response rule based on the amount of evidence collected over a longer time seems more appropriate. In this paper, we extend the idea of a simple race to continuous sampling of evidence in favor of a certain response category. The resulting Poisson random walk model is a TVA-based response time model in which categories are reported based on the amount of evidence obtained. We demonstrate that the model provides an excellent account for response time distributions obtained in speeded visual categorization tasks. The model is mathematically tractable and its parameters are well founded and easily interpretable. We also provide an extension of the Poisson random walk to any number of response alternatives. We tested the model in experiments with speeded and non-speeded binary responses and a speeded response task with multiple report categories. The Poisson random walk model agreed very well with the data. A thorough investigation of processing rates revealed that the perceptual categorizations described by the Poisson random walk were the same as those obtained from TVA. The Poisson random walk model could therefore provide a unifying account of attention and response times.

Keywords

Response time model; Attention: object based; Perceptual decision making, Math modeling

Public Significance Statement

A new model of response speed and response accuracy for recognizing similar looking visual objects is presented and tested in this article. The model is based on the assumption that observed behavior can be described by continuously categorizing visual objects until a pre-set response criterion is reached. In contrast to comparable sequential sampling models, the underlying theory is mathematically tractable and can be interpreted in cognitive terms, like speed of mental processing. Further, we present an interesting generalization of the new model that is not restricted to two perceptual categories.

Disclosures and Acknowledgements

The work is part of the Dynamical Systems Interdisciplinary Network, University of Copenhagen. Søren Kyllingsbæk was funded by The Danish Agency for Independent Research (Sapere Aude program), by the Danish Council for Independent Research, Social Sciences and by the Danish Council for Independent Research, Humanities. Steven Blurton was funded by the The Danish Agency for Independent Research (Award Number: 8018-00014B), by the Danish Council for Independent Research, Social Sciences and by the Danish Council for Independent Research, Humanities. These funding sources had no other role other than financial support.

Steven P. Blurton designed and conducted the experiments, analyzed the data, and wrote the manuscript. Søren Kyllingsbæk designed the experiments and wrote the manuscript. Carsten S. Nielsen wrote the manuscript. Claus Bundesen designed the experiments and wrote the manuscript. All authors have read and approved the final manuscript.

All authors declare no conflicts of interest.

We are constantly confronted with a vast amount of visual information. The visual system must efficiently process this stream of information and assign its limited resources to the most relevant objects in the visual field. The importance of this ability has led to many theories on visual attention over the last decades. One of these is the theory of visual attention (TVA) proposed by Bundesen and colleagues (Bundesen, 1987, 1990; Bundesen & Habekost, 2008; Bundesen, Habekost & Kyllingsbæk, 2007; Shibuya & Bundesen, 1988). With TVA it is possible to explain a variety of results in attention research, such as the effect of stimulus exposure duration on the accuracy of item report in experiments with partial and whole report paradigms (Busey & Townsend, 2001; Sperling, 1960; Townsend, 1981). The theory is a combination of a fixed-capacity independent race model (FIRM, Shibuya & Bundesen, 1988) and the biased choice model (Luce, 1963; Shepard, 1959). The TVA can explain empirical results from a variety of paradigms by two attention mechanisms: selection of features and selection of categories (Bundesen, 1990). These mechanisms are very general which relates TVA to a variety of other models (Logan, 1996; Logan, 2002, Logan & Gordon, 2001). Although predictions for response times (RT) have been derived from TVA for some experimental tasks, such as the spatial cueing task (Posner, Snyder, & Davidson, 1980; see Bundesen, 1990), the focus of research on TVA has so far been on the accuracy of report, especially in whole or partial report tasks.

At the same time, sophisticated RT models have been developed that predict whole distributions of correct and incorrect responses. These models often rest upon the assumption that information is accumulated over time until a pre-set criterion is reached. Observed variability in RT is primarily accounted for by assuming moment-to-moment variation during the decision process which is modelled by a stochastic process (Audley & Pike, 1965; Busmeyer & Townsend, 1993; Diederich, 1997; Link & Heath, 1975; Pike, 1973; Ratcliff, 1978; Smith, 2000; Vickers, 1970; but see Brown & Heathcote, 2005, 2008). Due to this moment-to-moment variation, the pre-set criterion is reached sooner in some trials and later in

other trials and the criterion associated with an incorrect response is occasionally reached sooner than the criterion associated with the correct response. Hence, these models can simultaneously account for error rates and response latency. The notion of moment-to-moment variation contrasts with that of TVA, in which noise is typically assumed to be negligible. The main reason for this discrepancy is that TVA is concerned with attentional processes during perception of highly discriminable stimuli, such as letters or digits, whereas speeded RT tasks usually involve speeded responses to highly confusable stimuli. In this paper, we propose a response time model based on central ideas of TVA and an extension of these to a mechanism for selecting a perceptual category based on evidence obtained over a longer time. The model inherits favorable properties of TVA and offers a detailed account of response times. Thus, the resulting model is a unified cognitive model of attention-based speeded RT tasks.

Theory of Visual Attention

The theory of visual attention (TVA) describes a fixed-capacity, parallel first-terminating (K -terminating) model. A central assumption in TVA is that attention can be conceived as choice; in terms of selecting categorizations of the type “item x belongs to category i ” (Bundesen, 1990; Logan & Gordon, 2001). The probability that the categorization “item x belongs to category i ” is the first to be selected (chosen) depends on its processing rate v (velocity):

$$P(x, i) = \frac{v(x, i)}{\sum_{y \in D} \sum_{j \in R} v(y, j)}$$

that is, its processing rate $v(x, i)$ relative to the sum of processing rates $v(y, j)$ across all categories j in the response set R and all items y in the stimulus display D . The processing rates are determined by selection among perceptual categories (response set) and selection among stimulus features (stimulus set). This is achieved by two attention mechanisms:

selection of categories (pigeonholing) and selection of object with particular features (filtering). For selection of categories it is assumed that processing rates are determined by perceptual evidence $\eta(x, i)$ and bias β_i :

$$v(x, i) = \eta(x, i)\beta_i \quad (1)$$

The multiplicative relationship in (1) means that high processing rates are only obtained if both perceptual evidence and bias of the categorization i are high—thus only categorizations with high perceptual evidence and high bias are likely to be chosen. This mechanism has been referred to as pigeonholing. The bias parameter is task-dependent and controlled by an executive (Logan & Gordon, 2001). If the task is to report letters on a display, for instance, perceptual bias would be high for letters, but low for other categories (e.g., digits). The perceptual evidence on the other hand is determined by the similarity between the stimulus and a template obtained by prior experience with items that belong to this perceptual category.

Selection among objects with different stimulus features is known as filtering and is incorporated into TVA by attention weights w_x :

$$w_x = \sum_{j \in S} \eta(x, j)\pi_j$$

where S is the stimulus set, $\eta(x, j)$ is the perceptual evidence that stimulus x belongs to category j and π_j are pertinence, or priority values. Filtering is a mechanism that is related to stimulus features. The difference from pigeonholing is that weights are determined by the stimulus set S rather than the response set R : The attention weight is conceived as the sum of perceptual evidence across all stimulus features, weighted by the pertinence value. Setting the pertinence value π_j high for features that are relevant for processing increases the attention weight w_x of all items x that belong to that feature category at the expense of all items that do

not belong to that category. This is because attention weights w_x are high only if both perceptual evidence and pertinence of the stimulus category are high.

If, for instance, the task is to report red letters in a display of red and blue letters (partial report paradigm), filtering becomes essential to successfully perform the task. The pertinence for *color red* would be high and pertinence for all other colors would be low. Red items then have high evidence strength for $\eta(x, \text{red})$, which is weighted by a high pertinence value for “red” so that the attention weights of red items are high. Blue items have high evidence strength for $\eta(x, \text{blue})$, which, however, is weighted by a low pertinence value for blue. On the other hand, blue items have low perceptual evidence for category “red”, that is, $\eta(x, \text{red})$ is low, which is weighted by a high pertinence value. Both these combinations therefore contribute only weakly to the attention weights of blue items, favoring the processing of red letters instead of blue. The combined effect of filtering and pigeonholing is incorporated into TVA by the rate equation:

$$v(x, i) = \eta(x, i) \beta_i \frac{w_x}{\sum_{z \in D} w_z} \quad (2)$$

where D is the set of items on the display. Attention weights are normalized so that each item in the visual field receives a relative weight (the weights of all items on the display sum to unity). This process is called filtering; if, in the above example, filtering is perfect, pertinence for color red $\pi_{\text{red}} > 0$ and pertinence is zero for all other colors (incl. π_{blue}). In effect, mostly red letters enter the race for being encoded into VSTM and therefore mostly red letters are reported. Filtering and pigeonholing can exert their influences independently because the categorical bias parameters β_i are set independently of filtering. Suppose that category i is relevant for action. Then, increasing the perceptual bias β_i of that category speeds up processing of the categorization “item x belongs to category i ” without affecting the attentional weights w_x .

An item x is encoded (selected), if and when any of the perceptual categorizations “item x belongs to category i ” is encoded in the visual short-term memory (VSTM). It is assumed that upon stimulus presentation, task relevant templates are activated in visual long-term memory and that bias and the attentional weights are determined for subsequent processing. At this later stage, filtering and pigeonholing occur according to the chosen pertinence values and bias parameters, since these determine the processing rates for the subsequent processing stage—the race for being encoded into VSTM. This is assumed to be a parallel processing stage with fixed processing capacity and independent racers. A racer in the context of the theory is a categorization of the type “item x belongs to category i ”.

Competition arises because VSTM has a limited storage capacity K so that only the first K items that finish processing can be stored in VSTM. It is assumed that a categorization is reported if and only if it becomes encoded into VSTM. Consider a visual display containing six letters that are all to be reported (whole report paradigm). Then, it is assumed that each of the six stimuli activates a template so that six racers enter the second processing stage to be encoded into VSTM. This parallel first-terminating architecture has proven suitable for highly discriminable visual objects. In this case, each stimulus activates only a single categorization and processing of all other categories is negligible.

Under these assumptions, performance is limited by two factors: First, the capacity limitation only allows for K items to be encoded; after all slots in VSTM have been filled up no further items can be encoded (and hence, reported). Empirical estimates obtained in previous studies (e.g., Bundesen, 1990; Shibuya & Bundesen, 1988) put the estimate of $K \approx 4$ —that is, about four items can be encoded in each trial. A more direct test of K was performed by Fisher (1984), who also reported a K around four items. The fixed storage capacity assumption introduces a negative dependence, but evidence for independent selection or positive dependence have also been reported (Townsend, 1981; Busey & Townsend, 2001).

The second limitation arises if the items presented in the visual display are only briefly presented and followed by a mask. The processing of an item is interrupted by mask onset; this may happen before a categorization for x is selected. Using a suitable mask effectively interrupts processing of the targets so that racers that have not finished processing at some fixed time will not finish processing at all. If a racer has not finished processing by time t —the stimulus-mask interval, or exposure duration—it will not be encoded in VSTM and not be reported. Clearly, the longer the interval, the higher the probability that the racers that entered the race for being encoded into VSTM have finished processing and—given that storage capacity is available in VSTM—are encoded in VSTM and subsequently reported. Hence, backward masks constrain performance, which is useful for testing predictions empirically (Sperling, 1960).

An assumption on the temporal properties of the race is necessary to obtain testable predictions from the model. For simplicity, consider a situation where only a single target is presented in the visual display (Bundesen & Harms, 1999). In that case, the limited storage capacity of the visual short-term memory is of no concern and filtering can be disregarded. If one assumes processing rates to be constant throughout the visual categorization process, the finishing times of the racers are exponentially distributed (e.g., Bundesen, 1990, p. 527). The assumption of constant processing rates yields a mathematically tractable version of the model. The exponential distribution has several desirable mathematical properties (e.g., Bundesen & Habekost, 2008; Luce, 1986; Townsend & Ashby, 1983) which simplify mathematical derivations considerably. Under the assumption of constant processing rates, the probability that the categorization “item x belongs to category i ” is made before time t is given by a (shifted) exponential distribution. A simple response rule for single stimulus categorization with this model is that the categorization “item x belongs to category i ” is reported if, and only if, this categorization is the *first* categorization of x to be selected—that is, the first to be encoded into VSTM. If the processing rates $v(x, i)$ are constant over time, the

probability that with n perceptual categories the categorization “item x belongs to category i ” is the first categorization of x to finish processing is

$$P(x, i) = \int_0^\infty v(x, i) \prod_{j=1}^n \exp[-v(x, j)t] dt$$

$$= \frac{v(x, i)}{\sum_{j=1}^n v(x, j)}$$

that is, the Shepard-Luce choice rule (Luce, 1963; Shepard, 1957).

If the target to be categorized is followed by a mask (backward masking), it is assumed that processing of the target is interrupted at mask onset. The probability of selecting a categorization for x then depends on the time between target and mask onset—denoted here as *exposure duration*. Specifically, the probability that item x has been categorized as belonging to category i within exposure duration t is given by the exponential cumulative distribution function $F(t | v)$ with rate parameter $v(x, i)$, the processing rate of the categorization “item x belongs to category i ”. Hence, $F(t | v)$ is the probability that upon presenting stimulus x , a perceptual categorization “ x belongs to category i ” has been encoded into VSTM by time t (Bundesen, 1990; Kyllingsbæk, 2006).

So far, we considered performance based on more or less perfectly distinguishable stimuli (such as letters or digits). Confusability between stimuli (and hence, perceptual categories) is low and selecting the correct perceptual category for stimulus x is only bounded by exposure duration (i.e., “misses”). In this situation, a simple race model consisting of a race between perceptual categorizations is appropriate. However, a race of this type is not always plausible (Bundesen & Harms, 1999, Kyllingsbæk, Markussen, Bundesen, 2012): In a more general situation with mutually confusable stimuli, several perceptual categorizations will be sufficiently activated in the visual long-term memory (VLTM). If more than one template is active, this inevitably leads to decision noise in the subsequent stage in which a

categorization is selected: Selecting the first categorization that finishes processing is not an optimal strategy since it will lead to frequent errors. If the first categorization is wrong, an incorrect perceptual category will be selected because in the race model there is no possibility to correct an incorrect first categorization (i.e., confusion). Hence, a simple race model seems implausible to account for erroneous categorizations (errors of commission/intrusion errors).

In the general situation with mutually confusable stimuli, categorization of these stimuli must involve a strategy that addresses the issue of more than one perceptual categorization being activated by a single visual object. One strategy is to assume that a single perceptual categorization might not be enough to confidently report a stimulus category. This implies that the selection of a perceptual category for x is not simply based on the first categorization to finish processing but on multiple tentative categorizations for x , obtained over longer time. Thus, a more general version of TVA for categorization of mutually confusable stimuli accounts for the problem that more than one perceptual categorization may be activated in VLTm by assuming that selection among those categories is based on multiple tentative categorizations obtained by sequential categorization of the stimulus to obtain as much information as necessary to select a perceptual category.

Such a generalization of TVA was provided by Kyllingsbæk et al. (2012), who employed a Poisson counter model to describe the categorization of highly confusable stimuli. In their model, categorization is described as a sequential sampling process at the end of which a perceptual category is selected based on a *what has most evidence* rule. This rule does not necessarily require an information criterion: in line with the instructions given, it was assumed that observers would collect as much information about the stimulus as possible during presentation of an item (as opposed to stopping after reaching a criterion). After stimulus presentation, the observer is assumed to carefully evaluate the accrued information and to faithfully report the perceptual category with most evidence. With this type of processing, the unbounded Poisson counter model predicts discriminability to monotonically

increase with the amount of available information. In case of two alternatives, represented by Poisson counters A and B with rates v_A and v_B , it holds that sensitivity d'

$$d' = \frac{E(X_A - X_B)}{\sqrt{\text{Var}(X_A - X_B)}} = \frac{(v_A - v_B)t}{\sqrt{(v_A + v_B)t}}$$

grows unboundedly with time.

Kyllingsbæk et al. (2012) demonstrated that the Poisson counter model is an excellent account of the accuracy of report in a pure accuracy task with mutually confusable stimuli. The model, however, is not readily applicable to response times because it lacks a response criterion.

Response time models

To obtain predictions for response time we used the evidence-accumulation framework—that is, we developed a model in which evidence is accumulated until a preset response criterion is reached (Ashby, 2000; Brown & Heathcote, 2005; Link, 1975; Nosofsky & Palmeri, 1997; Pike, 1973; Ratcliff, 1978, Smith, 2000; Smith & Vickers, 1988; Townsend & Ashby, 1983; Zhang & Bogacz, 2010; for models without this assumption, see e.g. Maddox & Ashby, 1996; Yellot, 1967). After having accumulated enough evidence, processing stops and an observable response is initiated (“threshold accrual halting models”, Townsend, Houpt, & Silbert, 2012). These RT models can be divided into two classes depending on whether they contain absolute or relative evidence criteria (Ratcliff & Smith, 2004; Smith & Ratcliff, 2004). Absolute evidence criteria are used for counter models (Pike, 1973; Townsend & Ashby, 1983) or accumulator models (e.g., Brown & Heathcote, 2005, 2008; Smith & Vickers, 1988; Vickers, 1970). The idea behind these models is that a response is initiated as soon as one of the counters or information accumulators reaches a pre-set evidence criterion. Depending on which process reaches its criterion first, a correct or incorrect response is given.

This contrasts with random walk (Link, 1975; Link & Heath, 1975) and diffusion models of RT (Ratcliff, 1978, 1981, 1988; Ratcliff & Rouder, 1998; Ratcliff & Tuerlinckx, 2002; Smith, 1995; Smith & Ratcliff, 2009) that contain relative response criteria. Here, detection does not only depend on the evidence accumulated in favor of one alternative, but also on the evidence obtained for the other alternatives. In random walk and diffusion models a response is initiated as soon as one alternative has accrued a critical amount of evidence *more* than any other alternative. The decisions and the time needed to reach that decision are modeled by a stochastic process that represents this difference in evidence. The response rule implicitly assumes that evidence for one perceptual category is taken as evidence against the other categories. Models of this class have been particularly successful in explaining response time distributions of correct responses and errors—including response accuracy¹—in a variety of tasks (Ratcliff & Smith, 2004), although recent work has cast doubt whether accumulator and random walk models are empirically distinguishable (Jones & Dzhafarov, 2014; Khodadadi & Townsend, 2015). For example, it is possible for an accumulator model to mimic predictions of a random walk model if the boundaries vary over time (Khodadadi & Townsend, 2015).

Random walk and diffusion models are not easily generalized to account for more than two response alternatives, however, more recently, more and more such generalizations have been developed (Kvam, 2019a, 2019b, Leite & Ratcliff, 2010; Ratcliff, 2018; Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007; Smith, 2016; Smith & Corbett, 2019). Counter or accumulator models, on the other hand, can easily be generalized to more than two alternatives, when the increments of each accumulator are assumed to be stochastically independent of the increments of the other accumulators (Brown & Heathcote, 2008, but see Brown & Heathcote, 2005; Usher & McClelland, 2001). If the increments are not assumed to

¹ We define response accuracy as the *conditional* probability of reporting the correct response category of the imperative stimulus, given that the stimulus elicited a response.

be independent, accumulator and random walk models can make identical predictions under some additional assumptions (Ditterich, 2006).

The first computational response time models were developed from the 1960s on (e.g., Audley & Pike, 1965; LaBerge, 1962; Laming, 1968; McGill, 1963; Pike, 1973; Stone, 1960). Early models of RT predicted mean RT (e.g., Link & Heath, 1975), whereas contemporary models predict complete distributions of correct and error RTs (Ratcliff, 1978; Brown & Heathcote, 2005). A general advantage of explicit RT models is that they provide a natural account of the speed-accuracy tradeoff. How well RT models account for the speed accuracy tradeoff is frequently tested by fitting these to data obtained in a condition in which response speed is stressed and one in which response accuracy is stressed (e.g., Ratcliff, Thapar & McKoon, 2001). The model then needs to account for data in both conditions with only the evidence criterion being variable across speed and accuracy (e.g., Ratcliff & Smith, 2004). In a more general interpretation of the speed accuracy tradeoff, explicit RT models like the diffusion model account for the tradeoff of response speed and accuracy by simultaneously accounting for response latency and response accuracy with parameters that independently account for response strategies or perceptual processes (Tanner & Swets, 1954).

Models based on the unbiased random walk, that is, a random walk between two boundaries equally apart from the starting point, are too restrictive in the sense that they predicted the same conditional distributions for error and correct RT distributions and, hence, mean RT (e.g., Ratcliff & Smith, 2004; Townsend & Ashby, 1983, p. 305). This prediction, however, is usually contradicted by empirical data: Dependent on the task and task instruction, error RTs can be faster and/or slower than correct responses.² It is usually the case that errors are faster than correct responses if the task is easy and response speed is stressed, but often slower than correct responses if the task is difficult and accuracy is stressed.

² Here and in the remainder of the article, references to fast or slow errors mean that the probability for fast or slow errors is higher than predicted by the unbiased random walk with its identical predictions for the conditional distributions of correct and error responses (conditioned on the decision).

Crossover patterns are also commonly found. Early models could accommodate such patterns only with implausible assumptions regarding their evidence criteria (see Townsend & Ashby, 1983, p. 279, for the Poisson counter model). Different patterns of correct and error RT distributions due to fast and slow errors can be accounted for by the assumption of trial-by-trial variation in several model parameters. For example, it is often assumed that the processing (or: drift) rate associated with a certain stimulus or stimulus feature varies from trial to trial (Brown & Heathcote, 2005; Ratcliff, 1978). The rate is higher in some trials and lower in others, enabling the model to account for slow errors. A similar assumption can be made for the starting point of the stochastic process: When the starting point is assumed to vary across trials, it may be relatively close to the criterion associated with the incorrect response in some trials, accounting for fast errors. The combination of trial-by-trial variation in starting point and drift rates provides enough flexibility for the models to account for various patterns that can be found in empirical RT distributions.

Finally, the non-decision time is sometimes considered a random variable that is independent of decision latency (Ratcliff & Tuerlinckx, 2002). If the residual latency is assumed to be constant, one can simply subtract it from the observed RTs; resulting in a shift of the predicted RT distributions along the time axis. If residual time is itself a random variable and model predictions on the level of RT distributions are required, a convolution of the decision time and non-decision time distributions is necessary (Luce, 1986, Ch. 3). For some models, analytic solutions to this convolution are available (e.g., Ashby & Townsend, 1980; Brown & Heathcote, 2008; Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011; Schwarz, 2001), whereas in mathematically more complex models this convolution is evaluated numerically (e.g., Ratcliff & Tuerlinckx, 2002). Verdonck and Tuerlinckx (2016) have suggested a different approach to deal with non-decision time: They proposed a method to factor out residual latency instead of specifying it explicitly in the model. In this case, only

the decision time needs to be accounted for by the explicit model, so that potential misspecifications of the non-decision time distribution are avoided.

In this paper, we present a random walk model based on TVA and the Poisson counter model. We argue that *speeded* visual identification can be represented by a Poisson random walk model—essentially a simple random walk with exponentially distributed interstep times (Bundesen & Harms, 1999). We assume that multiple tentative categorizations are obtained for x and that these represent evidence in favor of each perceptual category. The simple race model needs to be extended to specify how the additional categorizations are used and conflicting evidence is integrated to select a perceptual category. One way of obtaining such an extension is the Poisson counter model (Logan, 1996; Pike, 1973; Townsend & Ashby, 1983). In this model, a *what has most evidence* selection rule is assumed: Instead of reporting the first category that finishes processing, the reported category is based on multiple tentative categorizations representing evidence in favor of a certain perceptual category (Kyllingsbæk et al., 2012). In principle, it would be possible to introduce evidence thresholds to this Poisson counter model. The Poisson counter model, however, always predicts errors to be slower than correct responses (Audley & Pike, 1965) in a way that is often contradicted by data (Ratcliff & Smith, 2004).

Instead, we assume that evidence for one perceptual category is evidence against all other categories, leading to a random walk model. The equations involved in the simple random walk are well known (Feller, 1968; Cox & Miller, 1965); but the simple random walk has rarely been used as a model for cognition (but see Bundesen, 1982; Ratcliff, 1988). Previous applications of the random walk to modelling of response times included variants in which increments of the random walk were determined by the likelihood ratio of the two response alternatives, and thus, continuously distributed (Link & Heath, 1975). These increments were assumed to occur in discrete steps with a constant time between leading to a discrete time and continuous state space model of RT. The diffusion model (Ratcliff, 1978)

has the advantage over the random walk that it is continuous in time and thus avoids the need to relate discrete steps to time. The Poisson random walk model on the other hand is based on a simple random walk that is discrete both in time and space but the interstep times are continuously distributed so that the proposed model is discrete in space, but continuous in time.

A Poisson Random Walk Model

Based on the assumptions of TVA (Bundesen, 1990; Bundesen & Harms, 1999), we developed a response time model of speeded stimulus recognition/categorization of mutually confusable stimuli. Because it is based on TVA its terminology and parameters are fully compatible with TVA. The critical difference lies in the notion of confusability: In TVA it is assumed that observable responses are the result of a first-terminating parallel race model. This kind of processing may be termed *immediate perception* and it is a suitable account for the special case of stimulus displays that contain easily discriminable stimuli, such as letters or digits. However, we are often confronted with less discriminable visual stimuli. How can we expect to model categorization of such stimuli? It seems reasonable to assume that in this case the reported category is not based on only the first categorization that finishes processing, but rather on the number of tentative categorizations (*mediate perception*). Specifically, we assume that a conclusive categorization is made and its category reported once the number of tentative categorizations in favor of it exceeds the number of categorizations against it by some critical amount (Bundesen, 1982). Together with the assumption of exponentially distributed interstep times, these assumptions lead to the Poisson random walk model.

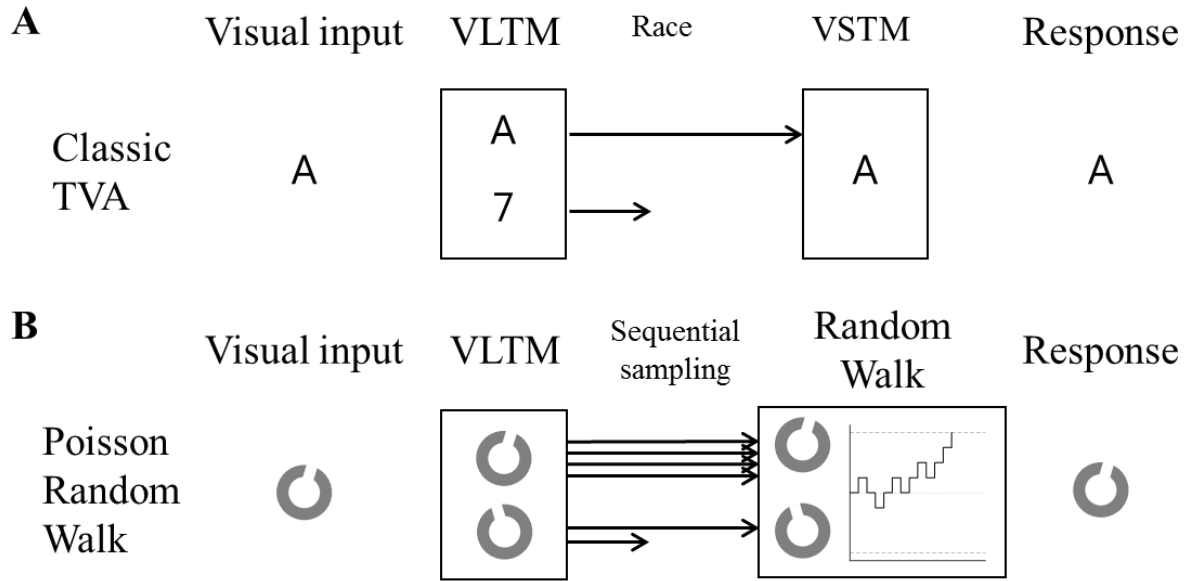


Figure 1. Schematic representation of the Poisson random walk as an extension of TVA. Visual input activates templates in visual long-term memory (VLTm). **A:** In TVA, items compete for being encoded in visual short-term memory (VSTM). This competition is conceived as a race between activated templates from VLTm with the first item to finish processing being the one reported (immediate perception). **B:** The Poisson random walk model is an extension in which evidence in favor of perceptual categories is continuously sampled until one alternative has a critical amount of counts more than the other (mediate perception). This selection can be modeled by a random walk.

As in TVA, we assume an early massively parallel stage in which the visual stimuli in the display activate visual object templates retrieved from VLTm. Matches against these templates lead to tentative categorizations. Formally, a tentative categorization is made if and when a racer finishes processing and we assume that the racers' finishing times are exponentially distributed (Bundesen & Harms, 1999; Nosofsky & Palmeri, 1997a). The probability that item x finishes processing—that is, is being categorized as belonging to category i , before time t is given by a Poisson distribution with rate vt . This sampling process proceeds until an evidence threshold is reached and the associated response is initiated.

For now, consider a task with two perceptual categories—that is, a task in which a set of stimuli x are to be categorized as belonging to one of two possible categories. After the initial activation of VLTm templates, the number of tentative categorizations for each

category i shall be represented by a counter $X_i(t)$. In case of two perceptual categories we define the response set $R = \{A, B\}$ so we may denote the counters associated with these responses $X_A(t)$ and $X_B(t)$. The counters $X_A(t)$ and $X_B(t)$ are assumed to be Poisson counters with rates $v_A(t)$ and $v_B(t)$, respectively. These rates may possibly depend on time t but for now we assume that these rates are constant. In other words, we assume a time-homogenous processing model with rates v_A and v_B . Let $D(t)$ be the difference between the two counters $X_A(t)$ and $X_B(t)$:

$$D(t) := X_A(t) - X_B(t)$$

Then, $D(t)$ can be modelled as a Poisson random walk between an upper absorbing state located at k_A and a lower absorbing state at $-k_B$. A categorization “item x belongs to category A” corresponds to a step towards the upper criterion k_A and a categorization “item x belongs to category B” to a step towards the lower criterion k_B . Let p be the probability of tentative categorization in favor of Category A (i.e., a step towards k_A), given there is a tentative categorization. Using the superposition principle of Poisson counters, $p = \frac{v_A}{v_A + v_B}$ (e.g., Townsend & Ashby, 1983, p. 273), and the probability for a step towards k_B is $1 - p = \frac{v_B}{v_A + v_B}$. These probabilities are independent of time; this holds also true for time-varying event rates (i.e., time-inhomogeneous processing), if the processing rates $v_A(t)$ and $v_B(t)$ are mutually proportional functions of time (Bundesen, 1990, 1993). We will assume that this is the case throughout the paper.

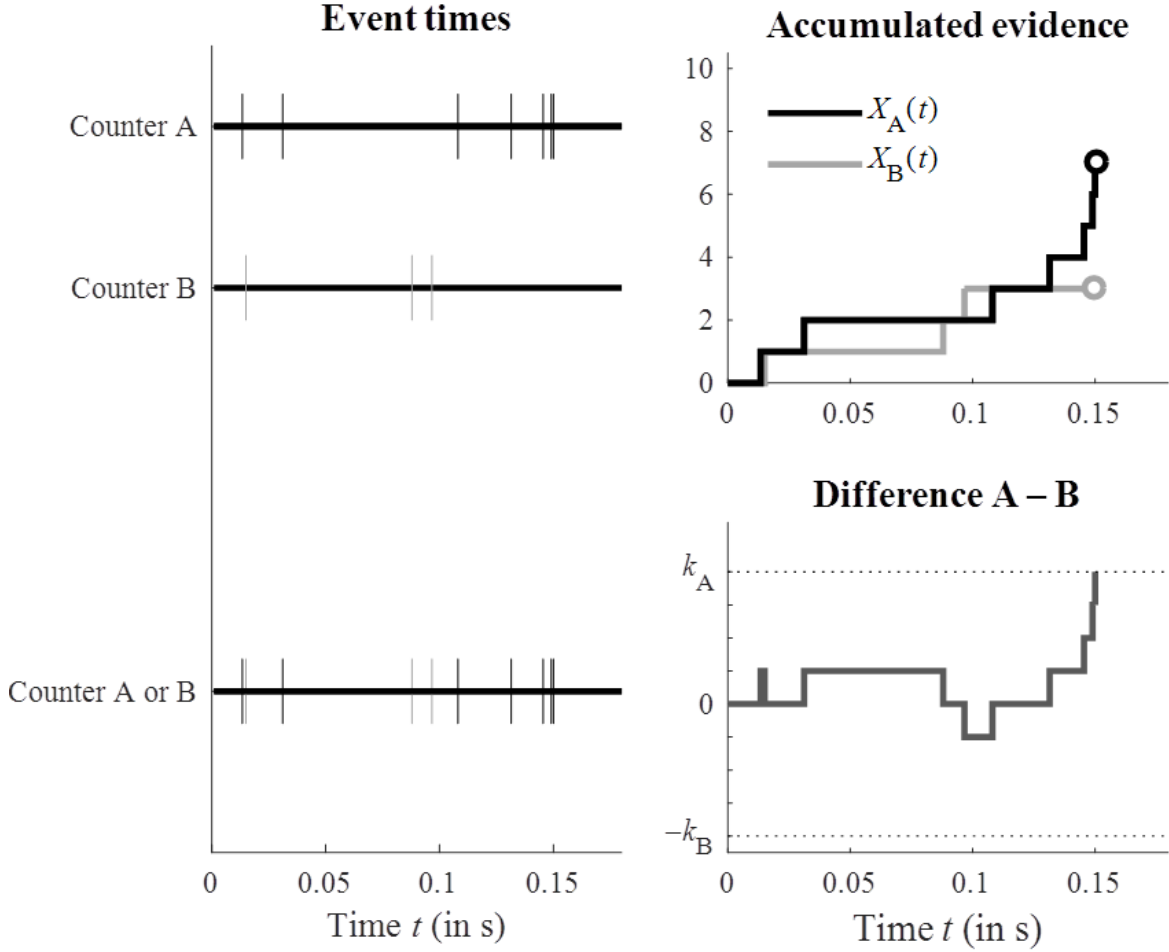


Figure 2. Random instance of a single trial. After stimulus onset, processing channels collect evidence in favor of perceptual category A or B in form of tentative categorizations. Accumulation of evidence for each category is modelled by a Poisson counter $X_i(t)$. A perceptual category is conclusively selected as soon as a category i (here: category A) has k_i tentative categorizations more than the other. This relative stopping rule leads to a random walk: $X_A(t) - X_B(t)$. Since the random walk is based on two Poisson processes, the event times at which the discrete process changes its states are obtained by superposition of the single Poisson processes. Parameter values used for this simulation were: $\nu_A = 60$ Hz, $\nu_B = 30$ Hz, and $k_A = k_B = 4$ categorizations.

The stopping criterion is based on the relative amount of information: Tentative categorizations are made until one category has accrued a critical number of tentative categorizations more than any other categories. This gives rise to a random walk between two absorbing barriers, which can be viewed as relative response criteria defined on the difference of the two counters (Figure 2). There are two perceptual categories and thus two evidence thresholds k_A and k_B , associated with selecting perceptual Categories A and B, respectively.

Define the random variable N_B as the number of steps needed for $D(t)$ to reach criterion k_B for the first time without having reached criterion k_A before. Without loss of generality, we set $a = k_A + k_B$, and introduce a starting point $z = k_B$, so that the upper criterion is placed at a and the lower criterion at 0 (Feller, 1968, p. 342). Let N_B be defined as the number of tentative categorizations needed for the random walk to reach the criterion for B for the first time and without having reached criterion for A before. The probability function of N_B , $f_B(n) = P(N_B = n)$ is known as the *first-passage time* probability function of a random walk between two absorbing barriers (e.g., Cox & Miller, 1965; Feller, 1968):

$$f_B(n | p, a, z) = \frac{2^n}{a} p^{\frac{n-z}{2}} (1-p)^{\frac{n+z}{2}} \sum_{j=1}^{a-1} \cos^{n-1} \left(\frac{\pi j}{a} \right) \sin \left(\frac{\pi j}{a} \right) \sin \left(\frac{\pi z j}{a} \right) \quad (3)$$

where $f_B(n | p, a, z)$ is the probability of absorption at 0, that is, selecting perceptual category B after exactly n steps, given probability p to move a step towards a , criterion separation $a = k_A + k_B$, and starting point $z = k_B$. By symmetry, $f_A(n | p, a, z) = f_B(n | 1-p, a, a-z)$ yields the probability of selecting perceptual category A after just n steps. As shown above, the probability parameter p is not a free parameter; since $D(t)$ is the difference of two Poisson processes with rates v_A and v_B , the probability p in (3) is $p = \frac{v_A}{v_A + v_B}$. Thus, the function $f_B \left(n \mid \frac{v_A}{v_A + v_B}, k_A + k_B, k_B \right)$ represents the probability for the count difference between two Poisson counters with rates v_A and v_B to reach k_B for the first time after exactly n steps and without having reached k_A before. However, for the sake of notional clarity, we use the a - z notation of the simple random walk throughout the paper.

Following from the assumption of Poisson processes and using the superposition principle of Poisson processes, the time between two events is exponentially distributed with a rate parameter equal to the sum of the Poisson rates: $C = v_A + v_B$. In cognitive terms, this sum of processing rates may be defined as visual processing capacity (Bundesen, 1990). Analogous to the Poisson counter model (e.g., Townsend & Ashby, 1983, p. 276), the

distribution function $F(t | n, C)$ of the waiting time T for n exponentially distributed processing times, each with mean $1/C$, to finish processing is:

$$F(t | n, C) = 1 - \exp(-Ct) \sum_{i=0}^{n-1} \frac{(Ct)^i}{i!} \quad (4)$$

Equation (4) is the cumulative distribution function of an Erlang distribution with shape parameter n and rate parameter C . The shape parameter n is the number of steps needed for the random walk to reach an evidence threshold and $C = v_A + v_B$, as defined above. As the number of tentative categorizations is a latent variable, the number of steps needed until absorption (n) is unknown; we therefore sum over all possible values of n to obtain

$$F_B(t | v_A, v_B, a, z) = \sum_{n=z}^{\infty} F(t | n, v_A + v_B) \cdot f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right) \quad (5)$$

with $f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right)$ given by Equation (3) and $F(t | n, v_A + v_B)$ by Equation (4).

Equation (5) is the cumulative distribution function of the time needed to select the perceptual category B . The distribution $F_B(t | v_A, v_B, a, z)$ is a weighted mixture of Erlang distributions with different shape parameter n , with the weights being determined by the first-passage time probability $f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right)$ of a simple random walk between two absorbing barriers

(Figure 1).

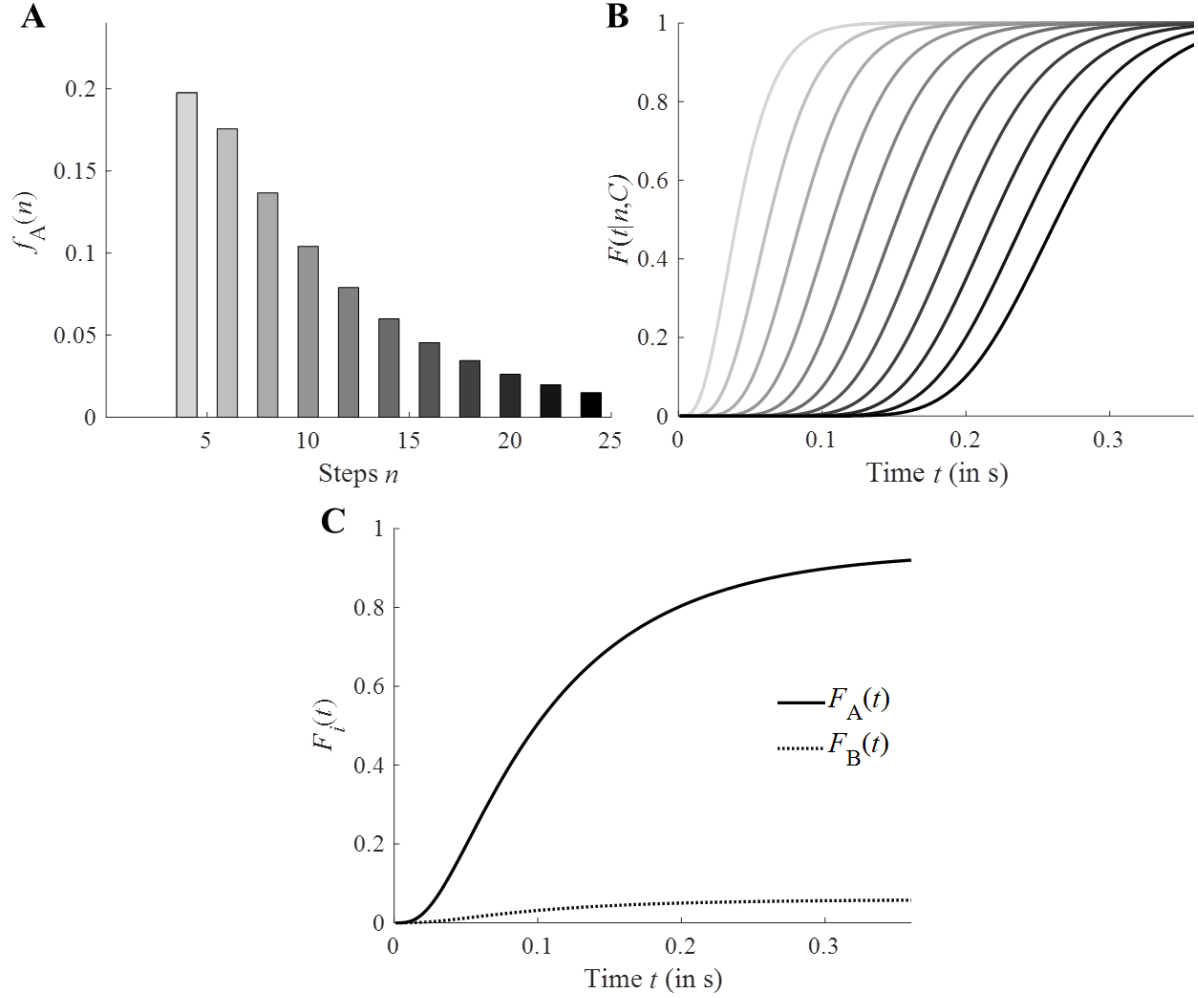


Figure 3. Components of the Poisson random walk. **A:** First-passage probability function $f_A(n | \nu_1/(\nu_1 + \nu_2), a, z)$ of a simple random walk between two absorbing barriers k_A and k_B . **B:** Waiting time distribution for n events with exponentially distributed latency and rate $C = \nu_1 + \nu_2$. The different grayscale values represent the shape parameter n of each distribution; the respective value for n can be obtained from Panel A, which uses the same grayscale coding. **C:** Cumulative distribution functions of reporting perceptual category A (solid line) or B (dashed line) in continuous time. Parameter values are: $\nu_A = 60$ Hz, $\nu_B = 30$ Hz, and $k_A = k_B = 4$ categorizations. Residual latency T_0 is not shown ($T_0 = 0$).

The model up to this point describes only the latency of selecting a categorization.

When fitting the model to response time data it is necessary to account for the latency of the remaining processes (“motor time”). The most common approach is to assume an additive RT model (e.g., Luce, 1986)—that is, $RT = T + T_0$. The simplest case, which is to assume a constant T_0 , merely requires a constant to be subtracted from the observed RT:

$$F_B(t|v_A, v_B, a, z, T_0) = \sum_{n=z}^{\infty} F(t - T_0|n, v_A + v_B) \cdot f_B\left(n \middle| \frac{v_A}{v_A + v_B}, a, z\right)$$

This concludes the derivation of the basic theoretical framework of a random walk model with exponentially distributed interstep times. In the following sections, we outline generalizations of the model with trial-by-trial variability in processing rates and starting point. Finally, the processing rates of the Poisson counters are assumed to depend on time t , yielding a time-inhomogeneous Poisson random walk model.

Trial-by-trial variation of model parameters

The basic version of the Poisson random walk model described in the previous section is a full model of speeded responses in two-alternative categorization—that is, it predicts RT distributions for correct and incorrect categorizations. However, as outlined above, random walk models of this type are usually too restrictive in the sense that in case of an unbiased random walk ($z = a/2$) they predict the conditional distributions of correct and incorrect categorization to be equal (Laming, 1968). Following Ratcliff (1978) and his extension of the diffusion model, we assume that processing rates in the Poisson counters are not constant over the course of an experiment but rather vary on a trial-by-trial basis. Noting that processing rates are defined only for positive reals ($v_A, v_B > 0$) we propose that the distribution of processing rates can be described by a gamma distribution. Accordingly, we assume

$$v_A \sim \text{gamma}(\theta_A, \xi)$$

$$v_B \sim \text{gamma}(\theta_B, \xi)$$

that is, the processing rates are gamma distributed with different shape parameters $\theta_A > 0$ and $\theta_B > 0$, and a common rate parameter $\xi > 0$. In the derivations, we will use the rate parameter notation of the gamma distribution, but we report the scale $1/\xi$ of the gamma

distribution throughout the article for better interpretability. This is because the scale parameter determines the variance of V as a multiple of its expectation: $\text{Var}(V) = E(V) \times 1/\xi$.

Let $G_B(t|\theta_A, \theta_B, \xi, a, z)$ represent the cumulative distribution function of selecting perceptual category B with gamma-distributed processing rates. Assuming independence between the processing rates of the Poisson processes, the problem is to find a solution to

$$G_B(t|\theta_A, \theta_B, \xi, a, z) = \int_0^\infty \int_0^\infty f(x|\theta_A, \xi) f(y|\theta_B, \xi) \sum_{n=z}^\infty F(t|n, x+y) \cdot f_B\left(n \middle| \frac{x}{x+y}, a, z\right) dx dy \quad (6)$$

where $f(x|\theta_A, \xi)$ and $f(y|\theta_B, \xi)$ represent the gamma probability density functions of the respective processing rates of the two Poisson processes and the sum corresponds to the first passage time distribution with constant rates (Eq. (5)).

It is possible to derive an analytical solution for this problem (see Appendix A):

$$G_B(t|\theta_A, \theta_B, \xi, a, z) = \sum_{n=z}^\infty \left[1 - \frac{\xi^{\theta_A+\theta_B}}{\Gamma(\theta_A+\theta_B)} \sum_{i=0}^{n-1} \frac{t^i \Gamma(i+\theta_A+\theta_B)}{i! (\xi+t)^{i+\theta_A+\theta_B}} \right] \frac{2^n}{a} \frac{B\left(\frac{n-z}{2} + \theta_A, \frac{n+z}{2} + \theta_B\right)}{B(\theta_A, \theta_B)} S(n|a, z)$$

with $S(n|a, z) = \sum_{j=1}^{a-1} \cos^{n-1}\left(\frac{\pi j}{a}\right) \sin\left(\frac{\pi j}{a}\right) \sin\left(\frac{\pi z j}{a}\right)$, the finite sum in (3), $\Gamma(\cdot)$ and $B(\cdot)$, the gamma and beta-function, respectively. Model parameters θ_A , θ_B , and ξ describe the distribution of rates v_A and v_B (see Equation (6)), parameter a is the separation of evidence thresholds ($a = k_A + k_B$) for two perceptual categories and $z = k_B$ is the starting point of the random walk.

Another common assumption is that the starting point z of the random walk also varies from trial to trial. In some trials, the process may start with a bias in favor of one or the other alternative—that is, the starting point is closer to an evidence criterion than in other trials. If

the starting point is close to the criterion associated with the incorrect responses in some trials, this may lead to fast errors, that is, errors that are faster than correct responses. Alternatively, evidence criteria may be modeled with trial-by-trial variation which has similar effects on model predictions (Ratcliff & Rouder, 1998). One motivation for including starting point variability is premature sampling—that is, it is assumed that observers start sampling (noise) before stimulus onset (Laming, 1968; Ratcliff & Rouder, 1998; Rouder, 1996). In the diffusion model, most frequently a uniform distribution is assumed for the starting point when fitting the model to data (Ratcliff & McKoon, 2008; Ratcliff, Thapar, & McKoon, 2001; Ratcliff & Tuerlinckx, 2002).

Together with trial-by-trial variation of processing rates (which accounts for errors that are slower than correct responses), starting point variation yields a model that can account for different patterns of correct and incorrect RT distributions. The exact starting point distribution is usually not important and often chosen for mathematical convenience (Jones & Dzhafarov, 2014). In the Poisson random walk with gamma-distributed processing rates, we considered the beta-binomial distribution to be a natural candidate for the starting point distribution as each increment in the random walk follows this distribution. The beta-binomial distribution has three parameters, parameter n denoting the number of trials and the two parameters of the underlying beta distribution (denoted here as z_A and z_B)

$$Z \sim BB(n_Z, z_A, z_B)$$

The width of the distribution (n_Z) was set to match the distance between the evidence thresholds minus two to avoid degenerations of the process (i.e., absorption at or before stimulus onset). That leaves two parameters (z_A, z_B) of the starting point distribution as free parameters which need to be estimated from the data. The beta-binomial distribution exhibits considerable flexibility; it can approximate the binomial distribution or the (discrete) uniform distribution, depending on its two shape parameters. For example, the uniform distribution is

obtained by setting $z_A = z_B = 1$. Generally, if $z_A \neq z_B$, the starting point distribution will be skewed—that is, biased in favor of one alternative (e.g., towards Category A, if $z_A > z_B$).

Despite its flexibility the beta-binomial distribution leads to a model that has only one parameter more than the model with a constant starting point.

In symbols, the probability distribution for starting point Z is

$$P(Z = z | n_z, z_A, z_B) = \binom{n_z}{z-1} \frac{B(z_A + z - 1; z_B + n_z - z + 1)}{B(z_A; z_B)}$$

for $z \geq 1$ and $z \leq n_z$, zero elsewhere. Parameters z_A and z_B are shape parameters of the beta binomial distribution, n_z is the fixed width ($n_z = a - 1$) of the distribution and $B(\cdot)$ is the beta-function.

Let $H_B(t | \theta_A, \theta_B, \xi, a, z_A, z_B)$ represent the first passage time distribution of the model with variability in processing rates and starting point. Since the state space of $D(t)$ is discrete,

$$H_B(t | \theta_A, \theta_B, \xi, a, z_A, z_B) = \sum_{z=1}^{a-1} G_B(t | \theta_A, \theta_B, \xi, a, z) \times P(Z = z | a - 1, z_A, z_B)$$

that is, a mixture distribution obtained by summation over all possible starting values z .

Time inhomogeneous processing

Up to this point the processing rates v_A and v_B are assumed to be constant *within* a trial. This restriction can be relaxed to allow for time inhomogeneous processing: In this case the processing rates $v_A(t)$ and $v_B(t)$ are zero at stimulus onset and gradually increase to their maximum levels (Figure 4). Such a response profile would be expected from sustained processing channels. In contrast, transient channels respond to changes in the environment; their activity is maximal shortly after stimulus onset with a sharp decline thereafter.

We followed Smith and Van Zandt (2000) in their differentiation between transient and sustained encoding channels (Smith, 1995, 1998). In principle, it is possible to include both sustained and transient channels in the model. Transient encoding filters are relevant for simple speeded RT tasks (detection tasks, see Burbeck & Luce, 1982) or pure accuracy tasks which often exhibit non-monotonic hazard rates (e.g., Petersen & Andersen, 2012). In contrast to simple speeded responses, response times of speeded categorization (discrimination) usually do not show a clear transient component—that is, even if transient channels contribute to solving the task, it is difficult to estimate transient contributions from empirical data. Hence, for the Poisson random walk model for speeded categorization we included only sustained encoding filters in the model. The processing rates $v_A(t)$ and $v_B(t)$ are assumed to be the output of those filters, serving as an approximation to the visual encoding stage. Predictions for time inhomogeneous processing are mathematically tractable for diffusion models (see Smith, 2000; Smith & Ratcliff, 2009) and even simpler expressions are obtained from Poisson models (Smith & Van Zandt, 2000).

The temporal profile of processing rates is assumed to be determined by sustained channels whose response profile is given by

$$\lambda_S(t) = v_S \times \left[1 - \exp(-\mu_S t) \sum_{j=0}^{n_S-1} \frac{(\mu_S t)^j}{j!} \right]$$

that is, a scaled gamma cdf, with shape parameters n_S ($n_S = 5$) and rate parameter μ_S . The scaling parameter v_S is the filter amplitude of those sustained channels; it corresponds to the constant processing rates in the homogeneous Poisson model.

The additional flexibility of time inhomogeneous processing comes with little mathematical complexity, if the ratio of processing rates of the Poisson counters is restricted to be constant (proportional hazards model, see Bundesen, 1993). This is the case if the temporal profile (i.e., μ_S and n_S) is identical for all counters and only the filter amplitude v_S

changes across counters. Under this restriction, the random walk part in (5) remains unchanged with time-varying processing rates under the assumption of proportional hazards, only the Erlang distribution in (5) needs to be adapted to account for time inhomogeneous processing.

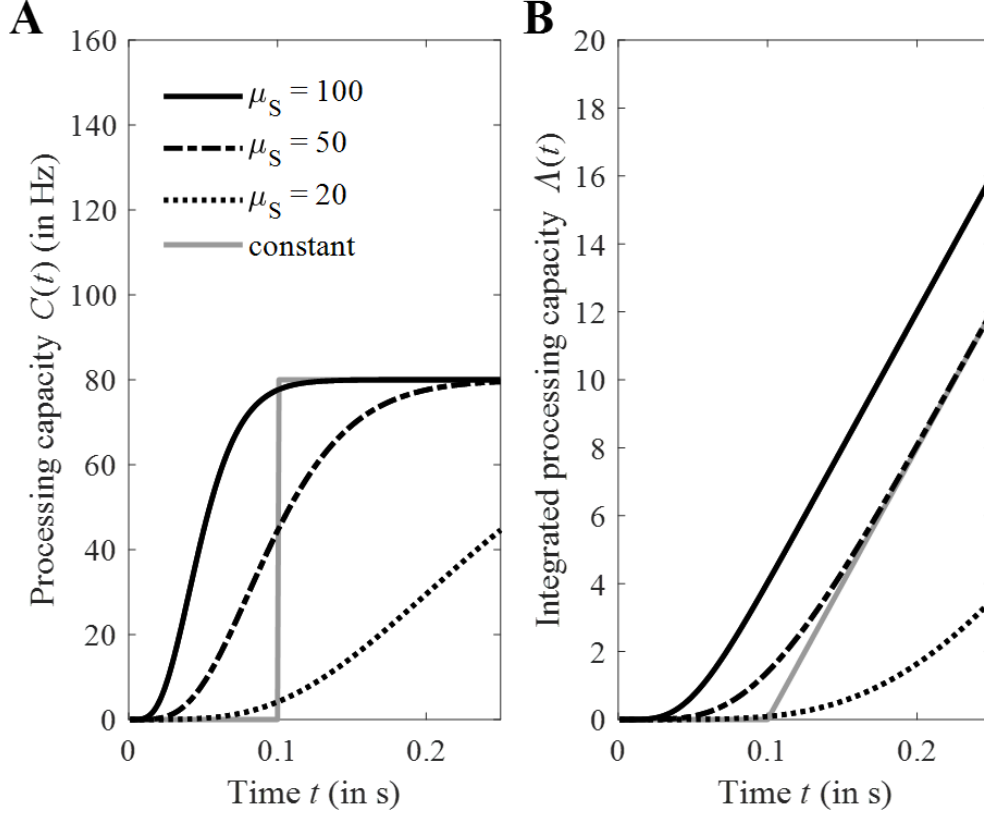


Figure 4. Effect of time inhomogeneous processing. **A:** Time-dependent processing capacity $C(t) = \sum v_i(t)$ (black lines, filter amplitudes: $v_1 = 60$ Hz; $v_2 = 20$ Hz) with varying μ_S ($n_S = 5$) compared to processing capacity with constant processing rates (gray line, $T_0 = 100$ ms). **B:** Integrated processing capacity $\Lambda(t)$ for time inhomogeneous (black lines) and time homogeneous (gray line) accumulation. The integrated event rate $\Lambda(t)$ is the expected number of categorizations up to time t . Except for a constant offset, the assumption of time inhomogeneous processing has very limited effects on integrated processing rates in the long term, but significant effects at very short times.

For time inhomogeneous processing, the integrated event rate $\Lambda(t) = Ct$ in (4) is exchanged by the generalized integrated event rate $\Lambda(t) = \int_0^t \lambda(s) ds$, where $v(s)$ may depend on time. Noting that the filter amplitudes are merely a scaling factor, let $\bar{\Lambda}(t) = \Lambda(t)/C$ be

the normalized integrated event rate. The model with time inhomogeneous processing is then obtained by modifying Equation (4) to

$$F[\bar{\Lambda}(t) | n, C] = 1 - \exp[-C\bar{\Lambda}(t)] \sum_{i=0}^{n-1} \frac{[C\bar{\Lambda}(t)]^i}{i!}$$

which is a temporally generalized Erlang distribution (Smith & Van Zandt, 2000). Explicit expressions for $\bar{\Lambda}(t)$ can be found in Appendix A.

The effect of sustained encoding filters is that the leading edge of the RT distributions can be modelled more accurately than with constant processing rates. This accounts for variability in the leading edge of the RT distributions (Smith & Ratcliff, 2009) and allows fitting the model with a constant residual “motor” time T_0 to response time data. Variability in the leading edge of the RT distribution can alternatively be accounted for by variability in non-decision time T_0 (Ratcliff & Tuerlinckx, 2002). For this approach, however, it is usually necessary to obtain predictions numerically as the required convolution of decision and non-decision time is not easily solved analytically. The assumption of time inhomogeneous processing, on the other hand, can be incorporated into the Poisson random walk with little additional mathematical complexity (see Appendix A for details). Another alternative that does not require an assumption about T_0 is the approach suggested by Verdonck and Tuerlinckx (2016). Factoring out non-decision time by the information provided by two or more RT distributions, avoids misspecifications of the non-decision time, or, in the case of temporally inhomogeneous processing, the time course of inhomogeneous processing.

Experiment 1

We conducted an experiment with a two-alternative categorization task to test the Poisson random walk model. The experiment included both a forced-choice, speeded response condition and a pure accuracy task with non-speeded responses and backward masking. In

both conditions, participants were instructed to judge the orientation of a Gabor patch. In the non-speeded condition, participants were presented with target stimuli followed by a mask and asked to respond as accurately as possible in an unforced choice task (pure accuracy task). In the speeded RT task, the same targets were presented without masking and participants were asked to respond as fast and accurately as possible in a speeded two-alternative forced choice task. The speeded RT data were used to estimate and test the Poisson random walk model. The non-speeded pure accuracy condition was then used to test if a Poisson counter model could accurately predict accuracy of report with the same parameters as in the speeded task.

Methods

Participants

We tested ten participants (mean age: 25.8 years, range 21–33 years, two left-handed) in an experiment that contained both a pure accuracy task and a speeded response time task (within-subject). Prior to the experiment, the participants gave their written informed consent. The study was approved by the local ethics board (Case no. 2014-01) and participants received a compensation of DKK 120 (15\$ approx.) for each hour of participation. All participants had normal or corrected to normal visual acuity.

Apparatus and stimuli

Target stimuli were full contrast (Michelson contrast $C_M = 0.998$) Gabor patches with six different orientations. We employed left-to-right gratings that were tilted from one degree up to three degrees to the left or right (i.e., -1 , -2 , -3 or $+1$, $+2$, $+3$ deg, respectively). The stimuli subtended 7 deg visual angle and were presented at the center of the visual display. In the pure accuracy task, the target stimulus was followed by a mask. The exposure duration (time between stimulus onset and mask onset) was systematically varied between 10 ms and

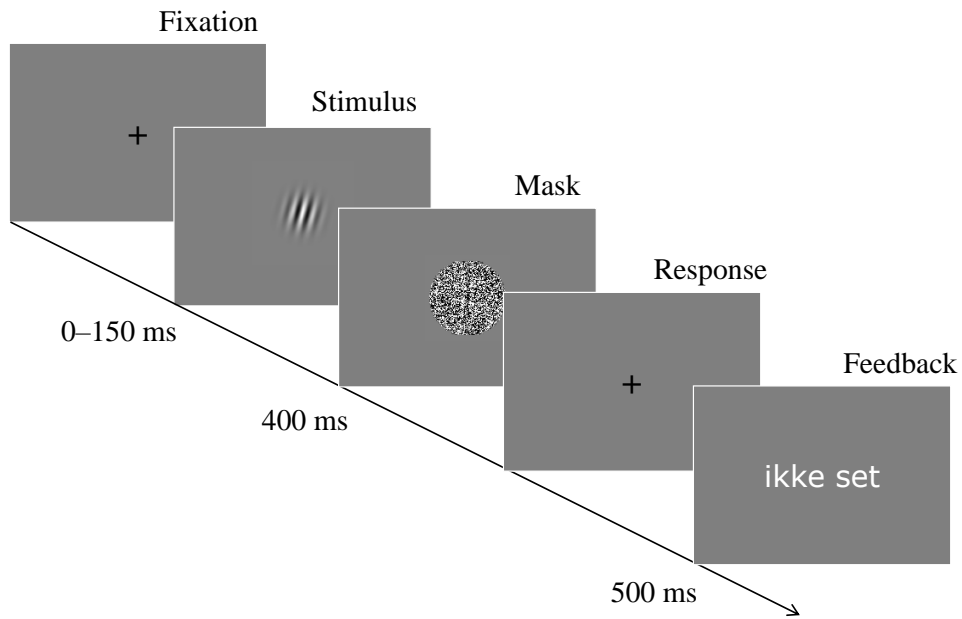
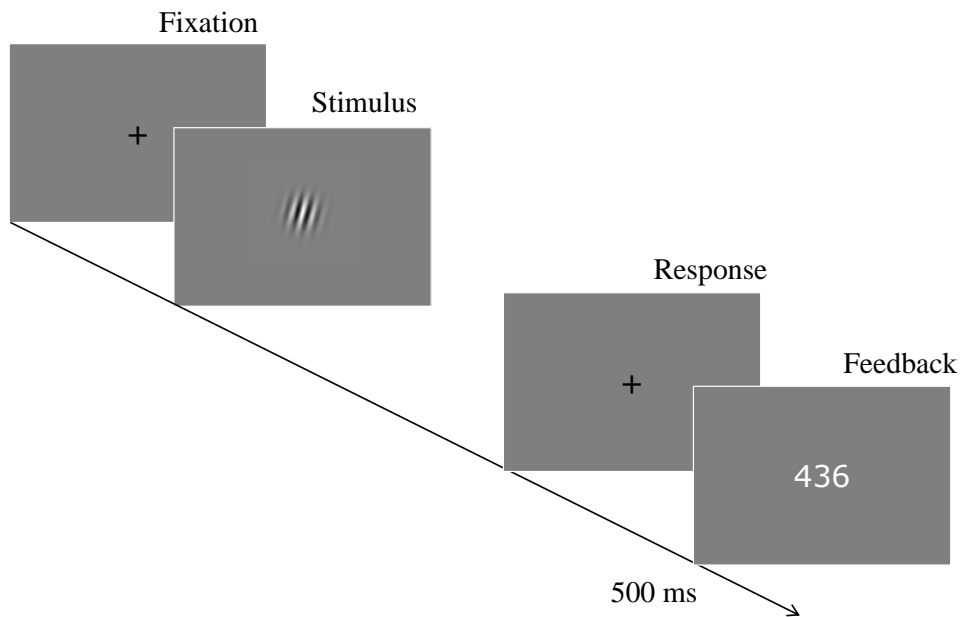
A**B**

Figure 5. Stimulus protocol in Experiment 1. **A:** *pure accuracy task.* Exposure duration of the stimulus was between 10 ms and 150 ms, 0 are catch trials. After the exposure duration, a mask (left-to-right grating with added noise) was displayed. Feedback was provided on the accuracy of the response (correct/incorrect). The decision was unforced; it was always possible to reject the trial (“not seen” feedback). **B:** *speeded response task.* There was no backward masking in this condition; the stimulus remained visible until a response was given. Feedback on response speed (RT in ms) was provided after the response.

150 ms, with exposure durations of 10, 20, 30, 40, 50, 60, 80, 100, and 150 ms. Additionally, we included a catch-trial condition with zero exposure duration to discourage blind guessing and to estimate the tendency to guess in absence of any information about the target.

The mask was a left-to-right grating (0 deg orientation) with random Gaussian noise added. The strength of the noise was based on the difference in grayscale values between the left and right oriented target in each difficulty condition (e.g., difference between 2 deg left and right Gabor for a 2 deg oriented Gabor patch, see Figure 5). This ensured that those parts of the screen that contained most task-relevant information were masked by most noise. The effectiveness of this mask layout was tested prior to the experiment (Eriksen, 1980). The mask turned out to be equally efficient for all orientations that were used in Experiment 1. The trials in the speeded response time task were similar, except that no backward masking was used, and the target stimulus was visible until the participant responded (Figure 5).

The experiment took place in a dimly lit room, the participants sat in front of a 22-inch CRT display (100 Hz screen refresh rate) with their head stabilized by a chin rest 60 cm in front of the screen. The stimulation protocol was implemented as a psychophysics toolbox script (Brainard, 1997) running under MATLAB (The MathWorks, Natick, MA). Different response boxes were used for the pure accuracy task and the speeded response time task. For the latter, we used a standard low-latency two-alternative response box connected to the serial port of the computer. For the pure accuracy task with unforced choice, we used a multiple-button response pad (Cedrus Corporation, San Pedro, CA).

Tasks and experimental procedure

In both the pure accuracy task and the speeded response time task the participant was instructed to judge the orientation of the Gabor patch (i.e., turned left or right) and press the left or right button on the response box accordingly. In the pure accuracy condition the participants were instructed to respond as accurately as possible, not to guess, and that

response latency would not matter. They should only respond if they were “*fairly certain*” to have seen a target and otherwise reject the decision by pressing a third button on the response box provided in this condition (unforced choice).

At the beginning of each trial the fixation cross was removed, and the target stimulus was displayed (except for catch trials, see below). In the pure accuracy task, the target was quickly followed by a mask which in turn was displayed for 400 ms. The fixation cross was presented again after mask offset and the program waited for the observer’s response. In both conditions we provided feedback after the response: in the pure accuracy task the feedback was given by changing the color of the fixation cross into green or red for 500 ms. If the participant rejected the decision, the text “not seen” (in Danish) was presented instead. In the speeded response time task, the measured response time of the previous trial was displayed at the center of the screen for 500 ms.

The trials with zero exposure duration were conceived as catch trials. The ‘correct’ response to a catch trial was to reject the decision. In the speeded response time task the fixation cross was removed for a period of one second without a target stimulus being displayed. If no response occurred during that interval, stimulus presentation resumed with the next trial. Between any two trials there was a variable inter trial interval, the length of which was drawn from an exponential distribution (mean: 1.5 sec) with an additional constant interval (0.5 sec).

We used a block-wise presentation of pure accuracy trials and speeded response time trials. Each participant was tested in 10 blocks of each response task. Each block comprised 360 target trials plus an additional 36 catch-trials in both response tasks. We collected 600 replications in each of the six orientation conditions in the speeded RT task. Due to the variation in exposure duration in the pure accuracy task, there were 60 different conditions in that task. Each of the 60 conditions was tested 60 times over the course of the experiment—that is, we ran 3,600 trials in the pure accuracy task also.

Due to the length of the experimental testing procedure, it was split into three separate sessions conducted on three different days. Before the first block in each condition, we ran a testing sequence for practice. This sequence contained a reduced number of trials. Training for the pure accuracy task contained only the five longest exposure durations. The sequence was repeated until the participant achieved a mean RT of about 500 ms (speeded RT task) or a mean accuracy of above 70% (pure accuracy task). The data of all practice sequences were discarded. Each of the three sessions began with either pure accuracy or speeded response time task and then alternating until the end of the session. The task in the first block of each session was counterbalanced between and within participants (i.e., over sessions).

Data analysis

The model was implemented in MATLAB and fitted to data using a simplex algorithm (Nelder & Mead, 1965). The deviation between model predictions and data was assessed by a log-likelihood function. The response time data of the speeded response task was fitted using the quantile maximum likelihood function (Heathcote, Brown, & Mewhort, 2002). We used the 10%, 30%, 50%, 70%, and 90% distribution quantiles (e.g., Smith, Ratcliff, & Wolfgang, 2004) of the predicted and empirical cumulative response time distributions of correct and incorrect RTs. First, we conducted a separate model fit to the data obtained in the speeded response time task to assess the ability of the model to explain the shape of observed correct and error RT distributions. Subsequently, we fitted a combined model to the whole dataset under the restriction that the parameters determining the processing rates are the same in the speeded RT and pure accuracy task. More details on the fitting procedure can be found in Appendix C.

No data cleaning was performed for either the response time data or pure accuracy data. We entered RTs to all 3600 targets into analysis to estimate RT distribution quantiles and accuracy of report. For the RT distributions, trials were separated for stimulus condition

(600 RTs each) and RT quantiles were estimated from those 600 RTs. For the quantile probability plots, we estimated the quantiles from the conditional distributions of correct and error responses. The estimates of the conditional distributions were then averaged (“vincentization”, Ratcliff, 1979) across participants and across left and right responses. We used a bootstrap procedure to estimate confidence intervals of the quantiles. Experimental conditions were specified as fixed effects and participants as random effects—that is, in each iteration we took a random sample from our participants ($N = 10$) and then random samples from their 600 observed responses in each condition. This procedure was repeated 10,000 times to obtain a group estimate of the quantiles and a 95% confidence interval by calculating the 2.5% and 97.5% quantile of the resulting distributions of quantiles. The same approach was used to obtain group estimates of pure accuracy data. The confidence intervals were used to assess the goodness of fit of the model: if the model predictions lie within the 95% confidence interval of all data points, the model fit was deemed reasonably good. In addition to this assessment, we assessed the goodness-of-fit based on a pseudo- R^2 (Nagelkerke, 1991), which was calculated as

$$R^2 = \frac{1 - \exp(LL_0 - LL_1)^{2/n}}{1 - \exp(LL_0)^{2/n}}$$

in which LL_0 is the log-likelihood of the restricted model, LL_1 is the log-likelihood of a less restricted model, and n is the number of trials. In cases where there was no second model to compare to, we set the LL_0 as the log-likelihood of the single model and LL_1 as the likelihood of a hypothetical model that perfectly matches the data (“achievable likelihood”). We consider this a useful descriptive measure of the goodness-of-fit, reflecting the percentage of variation explained by the model. Being a descriptive measure for effect size, it has no unit and the additional advantage of being (approximately) sample-size independent (Nagelkerke, 1991).

For the statistical analysis of the behavioral data (i.e., mean RT, percent correct), we estimated mixed models with *eccentricity* (3, 2, and 1 deg), *side* (left/right), and *exposure duration* (in the pure accuracy condition) as fixed effects and *participant* as random effect. The percentage of correct responses was analyzed with a mixed-effect logistic regression model. Contrary to the Poisson model fit, we treated trial rejections in the pure accuracy condition as incorrect in this analysis and excluded all catch trials. For the speeded RT condition, we estimated the mean RT based on the Kaplan-Meier estimate of the survival function (Blurton, Greenlee, & Gondan, 2015; Koch et al., 2013) and analyzed those with a linear mixed-effects model. All statistical analyses were performed with R statistical language (R Core Team, 2018).

Results

Anticipatory responses and guesses

We did not observe responses to catch trials in the speeded response time task except in one participant (one response). Catch trials in the *speeded response task* were therefore not further analyzed. In the *pure accuracy task*, however, guessing was quite frequent despite the instruction given to the participants. In most participants, the observed guessing rates ranged between 5% and 15%. However, we also observed a response frequency in catch trials as high as 90% in one participant. In the *pure accuracy condition*, response frequencies to catch trials were therefore explicitly modelled (as $ED = 0$ condition).

Pure accuracy task

The main effects of target orientation ($\Delta LL = -21.31$, $\chi^2(1) = 42.60$, $p < .001$), confusability ($\Delta LL = -405.4$, $\chi^2(2) = 810.7$, $p < .001$), exposure duration ($\Delta LL = -5,177.6$, $\chi^2(1) = 10,356$, $p < .001$) as well as the interaction between confusability and exposure duration ($\Delta LL = -342.4$, $\chi^2(2) = 684.7$, $p < .001$) were significant. No higher order interaction with target

orientation was significant (all $\Delta LL < 1$, n.s.). As expected, the proportion of correct responses increased with exposure duration (odds ratio [OR] = 1.47³, $Z = 79.835$, $p < .001$). The accuracy of low confusability targets was significantly higher than that of targets with intermediate confusability (OR = 1.31, $Z = 8.31$, $p < .001$) and high confusability (OR = 2.37, $Z = 27.28$, $p < .001$). Accuracy in the high confusability condition was also significantly lower than in the medium confusability condition (OR = 0.55, $Z = -19.30$, $p < .001$). Finally, the participants responded to targets oriented to the right with significantly higher accuracy than to targets oriented to the left (OR = 1.18, $Z = 6.47$, $p < .001$). The two-way interaction between confusability and exposure duration was such that accuracy increased with exposure duration, but more so with lower confusability.

Speeded RT task

The error rates in the speeded RT task ranged between 0.1% and 9% in the easiest conditions (3 deg targets), between 1.5% and 11% in the intermediate condition (2 deg targets) and between 5% and 25% in the most difficult conditions (1 deg targets). The overall effect of both confusability ($\Delta LL = -414.2$, $\chi^2(2) = 828.5$, $p < .001$) and target direction ($\Delta LL = -64.62$, $\chi^2(1) = 128.5$, $p < .001$) on response accuracy was significant, and so was their interaction ($\Delta LL = -10.32$, $\chi^2(2) = 20.64$, $p < .001$). We observed a pattern similar to that in the pure accuracy task: compared to highest confusability (1 deg), the probability of correct responses was significantly higher in 2 deg targets (OR = 2.91, $Z = 20.06$, $p < .001$) and 3 deg targets (OR = 4.29, $Z = 23.96$, $p < .001$). Additionally, response accuracy in 3 deg was significantly higher than in 2 deg (OR = 1.47, $Z = 5.52$, $p < .001$). As in the pure accuracy task, response accuracy was significantly higher for targets oriented to the right (OR = 1.64, $Z = 11.04$, $p < .001$). The interaction becomes evident when evaluating the effect of orientation on accuracy for the different levels of confusability: The difference in accuracy of

³ Exposure duration coded in a hundredth of a second (i.e., a single frame) for better interpretability

responses to rightward compared to leftward tilted Gabor patches was largest with highest confusability (OR = 1.91), smaller with medium confusability (OR = 1.59), and smallest with low confusability (OR = 1.10).

In addition to the error rates, increasing confusability (i.e., task difficulty) exerted the usual effects on RT distributions: both correct responses and errors became slower with increasing confusability. In relative terms, we observed some particularly slow errors (compared to correct response RT distribution) in the most difficult conditions which can be seen in the longer tail of the error RT distribution of the difficult conditions (i.e. the two central distributions in Figure 6). In the two easier conditions, errors were often faster than correct responses, shifting the leading edge of the error distributions towards faster responses relative to the correct RT distributions (i.e., two outer distributions in each of the left and right panel of Figure 6).

Regarding mean RT, we observed significant effects of confusability ($F_{2,45} = 38.807$, $p < .001$) and orientation ($F_{1,45} = 11.866$, $p = .001$); their interaction was not significant ($F_{2,45} = 2.99$, $p = .060$). The participants were significantly slower to respond correctly in the 1 deg condition compared to the 2 deg condition ($t_{48} = 5.932$, $p < .001$) and compared to the 3 deg condition ($t_{48} = 7.216$, $p < .001$). These differences were 61.2 ms (95%-CI: 41.0 to 81.4 ms) and 74.5 ms (56.3 to 92.7 ms), respectively. Responses in the 3 deg condition were on average 13.2 ms (−6.9 to 33.5 ms) faster than responses in the 2 deg condition; this difference was not significant ($t_{48} = -1.285$, $p = .205$). Responses to targets oriented to the right were significantly faster than responses to targets oriented to the left ($t_{49} = -2.126$, $p = .039$), this differences was on average 25.4 ms (95%-CI 1.8 to 49.0 ms).

Fit of the Poisson random walk model to response time data

We fitted the Poisson random walk model to the response time data of the 2×3 (left/right \times Gabor patch orientation) experimental conditions, allowing for trial-by-trial

variability in processing rates and starting point. We conceived the processing rates of categorization “item x belongs to category i ” to be composed of perceptual evidence $\eta(x, i)$ and category bias β_i . Since the processing rates were modelled as gamma-distributed to account for trial-by-trial variability, the multiplicative composition of perceptual evidence and bias was assumed to hold on the level of the mean processing rate:

$$\bar{v}(x, i) = \bar{\beta}_i \bar{\eta}(x, i)$$

where bar parameters denote expectations. No attempt was made to distinguish different sources of variability of processing rates—that is, the exact distributions of β_i and $\eta(x, i)$ were not specified. Instead, we modelled the trial-by-trial variation of processing rates as the combined effect of variation in bias and variation in perceptual evidence. Under this assumption the processing rates are approximated by a gamma distribution with mean

$$\bar{v}(x, i) = \frac{\eta_\theta(x, i) \beta_\theta(i)}{\xi}, \quad (7)$$

where the product $\eta_\theta(x, i) \beta_\theta(i)$ is the shape parameter and $1/\xi$ is the common scale parameter of the gamma distribution. To obtain a reasonably constraint model, we restricted the perceptual evidence for correct ($\eta_\theta(x, C)$) and incorrect categorizations ($\eta_\theta(x, I)$) to be the same for both perceptual categories within each difficulty level and to depend only on the confusability of the stimulus (i.e., difficulty). One of the bias parameters was a scaling factor: The higher of the two bias parameters was fixed at unity: $\max_i \beta(i) = 1$ and the lower of the two was a free parameter; this was done individually for each participant to ensure that the estimates for eta are approximately on the same scale. Accumulation of perceptual evidence was modelled as a time inhomogeneous process by including sustained encoding channel activation encoding channels (μ_S) into the model. The common scale parameter of the distribution of processing rates ($1/\xi$), evidence thresholds ($k = k_A = k_B$), starting point

variation (z_A and z_B) as well as residual time (T_0) were assumed to be invariant across conditions. Thus, we used a total of 13 free parameters to fit five quantiles of each correct and incorrect response time distribution in the six conditions (i.e., 12 RT distributions) as well as the overall probability for a correct response in each condition. All fits were obtained for each individual and then averaged across participants. Table 1 contains the estimated model parameters averaged across all participants.

Table 1

Group average ($\pm SD$) processing rates and parameter estimates of the Poisson random walk RT model

Parameter			Group average ($\pm SD$)	
			Correct categorization	Incorrect categorization
$\bar{v}(x, i)$ (in Hz)	Stimulus orientation	3 deg left	70.7 (± 25.7)	17.3 (± 14.1)
		2 deg left	64.5 (± 24.3)	18.9 (± 14.3)
		1 deg left	52.9 (± 22.1)	25.4 (± 16.9)
		3 deg right	77.8 (± 30.5)	15.5 (± 12.5)
		2 deg right	71.2 (± 29.2)	16.8 (± 12.5)
		1 deg right	58.4 (± 26.1)	22.7 (± 14.7)
$\eta_\theta(x, i)$	Stimulus orientation	3 deg	304 (± 877)	117 (± 353)
		2 deg	290 (± 837)	121 (± 363)
		1 deg	247 (± 716)	144 (± 429)
$\beta_\theta(i)$	Perceptual category	“left”	0.92 (± 0.07)	
		“right”	1.00 (± 0.01)	
$1/\xi$			3.51 (± 3.23)	
k			4.09 (± 1.12)	
z_A			5.57 (± 9.59)	
z_B			5.63 (± 9.62)	
μ_S			51.9 (± 16.0)	
T_0 (in s)			0.239 (± 0.020)	

Note—Bar parameters \bar{v} denotes mean processing rates. Processing rates are gamma distributed with mean $\beta_\theta(i)\eta_\theta(x, i)/\xi$ and scale $1/\xi$. The higher of the two perceptual bias parameters $\beta_\theta(i)$ was a scaling parameter (determined on an individual basis). Criterion parameter k represents placement of evidence thresholds ($k = k_A = k_B$). The starting point distribution is beta-binomial with parameters z_A and z_B . Sustained encoding channels activity is determined by μ_S and T_0 is residual latency. Mean processing rates are not free parameters but calculated according to Equation (7) with the parameters given in the table. Thus, the model had a total of 13 free parameters. The standard deviation of parameters (in parentheses) was obtained across participants.

The model fitted well to data obtained in the speeded RT task (Figure 6). The agreement between model and empirical data was excellent: there was no systematic deviation between model predictions and data and all model predictions lay within the 95% bootstrap confidence intervals. Across all participants, the model accounts on average for 97.6% (95.7 to 98.7%) of variation in the data.

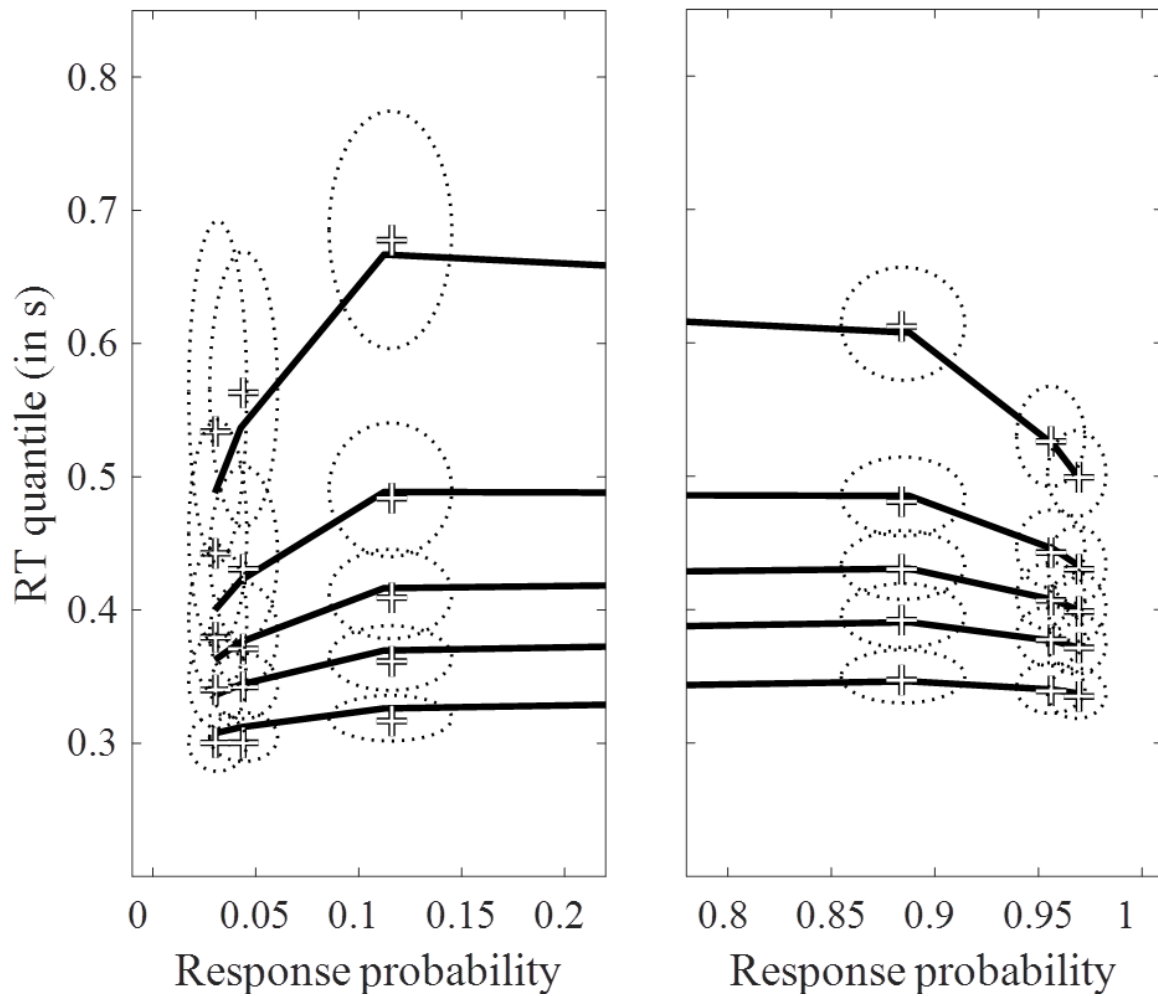


Figure 6. Quantile probability plot of the observed RT distributions (crosses) represented by their 10%, 30%, 50% (median), 70%, and 90% quantiles (from bottom to top). Model predictions of the Poisson random walk are represented by unmarked points connected with solid lines. The ellipses represent 95% confidence intervals of response accuracy and RT quantiles (as determined by bootstrap). Model parameter estimates for the displayed model predictions are summarized in Table 1.

The agreement between theory and data is particularly close for the distributions of correct responses (right panel in Figure 6) but also the distributions of incorrect responses were adequately predicted by the model. The model could well predict both fast errors (leading edge of the error RT distributions) and slow errors (median and tail of the error RT distributions). The probabilities of correct (incorrect) responses were also well predicted.

Perceptual evidence varied with confusability; the higher the confusability, the lower the perceptual evidence for the correct categorization and the higher the perceptual evidence

for the incorrect perceptual category. The resulting mean processing rates ranged between 55 Hz and about 80 Hz, visual processing capacity C between 80 Hz and 95 Hz. A mixed-effects linear model showed that the processing rates were significantly different across confusability condition for both the correct categorization ($F_{2,18} = 59.88, p < .001$) and the incorrect categorization ($F_{2,18} = 25.30, p < .001$).

The processing rate for the correct categorization in the lowest confusability condition (3 deg) was significantly higher than both in the 2 deg condition ($t_{18} = 3.72, p = .002$) and in the 1 deg condition ($t_{18} = 10.77, p < .001$). Additionally, the processing rates in the 2 deg condition were significantly higher than in the 1 deg condition ($t_{18} = 7.06, p < .001$). The difference between 3 deg and 2 deg was 6.4 Hz (95%-CI: 3.0 to 9.8 Hz), between 3 deg and 1 deg 18.6 Hz (15.2 to 22.0 Hz), and 12.2 Hz (8.8 to 15.6 Hz) between 2 and 1 deg conditions. The processing rates for the *incorrect* categorization were significantly higher in the highest confusability condition than those in the 2 deg condition ($t_{18} = 5.44, p < .001$) and the 3 deg condition ($t_{18} = 6.69, p < .001$). The difference between these conditions were 6.2 Hz (95%-CI: 4.0 to 8.5 Hz) and 7.6 Hz (5.4 to 9.9 Hz), respectively. The difference in processing rates of the two less confusable stimulus conditions was smaller (1.4 Hz, -0.8 to 3.7 Hz) and not significant ($t_{18} = 1.25, p = .229$). There was a perceptual bias in favor of the “rightwards” categorizations ($\beta_{\theta}(\text{“left”}) = 0.92, \beta_{\theta}(\text{“right”}) = 1.00$); all except one participant had a perceptual bias in favor of “rightwards” categorizations. Consequently, the processing rates for “rightwards” categorizations were significantly higher than those for “leftwards” categorizations ($t_9 = 3.16, p = .012$). The average bias across all confusability levels and correctness of the categorization was 7.1 Hz (2.5 to 11.7 Hz).

In sum, the processing rate estimates are comparable to those obtained in previous applications of TVA (e.g., Habekost, Petersen, & Vangkilde, 2014; Petersen, Kyllingsbæk, & Bundesen, 2012). The processing rate estimates were obtained from a model in which perceptual evidence ($\eta_{\theta}(x, i)$) depended on confusability, but not on the response category. To

test the appropriateness of this restriction, we compared the model to a “full” model, that is, one in which processing rates may depend on both confusability and perceptual category. A nested model test yielded significant results for half of our participants. Given the high number of trial repetitions, this is fairly weak evidence in favor of the full model. Indeed, in terms of pseudo- R^2 , the addition of those five extra parameters did not explain more than 1% of variation in any participant (0.07 to 0.95%).

Perceptual bias was slightly higher for the perceptual category “oriented rightwards” ($\beta_\theta(\text{“left”}) = 0.92$; $\beta_\theta(\text{“right”}) = 1.0$). This bias was significant ($t_9 = 3.34$, $p = .009$). The average evidence threshold was $k = 4.09$, meaning that the number of excess categorizations needed in favor of a perceptual category to be reported was about four categorizations for both categories. Starting point variability was quite high, as suggested by the presence of many fast errors resulting in a shift in the error RT distributions relative to correct RT distributions (Figure 6). In all but one participant ($z_A = 167.6$, $z_B = 168.4$) we obtained approximately a uniform starting point distribution ($z_A = z_B = 1$). The sustained channels responded quite slowly after stimulus onset; a sustained channel parameter μ_S of about 50 means that after 100 ms, only half of the maximum processing capacity was attained (see Figure 4). Finally, the latency of residual processes was 239 ms.

The effect of a continuous increase in processing speed on the model predictions are summarized in Figure 7 in terms of the estimated hazard rates (estimated from RT data): The time inhomogeneous model gives a noticeably better approximation of the leading edge of the observed RT distributions than the time homogeneous model. Although this difference may seem small, this part of the hazard function is estimated with highest precision; with increasing t , the hazard rates become more and more noisy since they are estimated from fewer data.

Adding trial-by-trial variation in non-decision time to the model with time inhomogeneous processing did not substantially improve the model fits. The average

estimated variation in non-decision time (uniform distribution with width s_t) was 23 ms (SD: ± 40 ms). This only changed when we restricted the model to time homogeneous processing: Then, trial-by-trial variation was 95 ms on average (± 27 ms). Consequently, variations in the leading edge of RT distributions were well accounted for by encoding processes, whereas the variation of non-decision time was rather low. The validity of this interpretation may be difficult to assess since both assumptions affect the leading edge of the RT distribution, so it may be impossible to divide these two sources of variation.

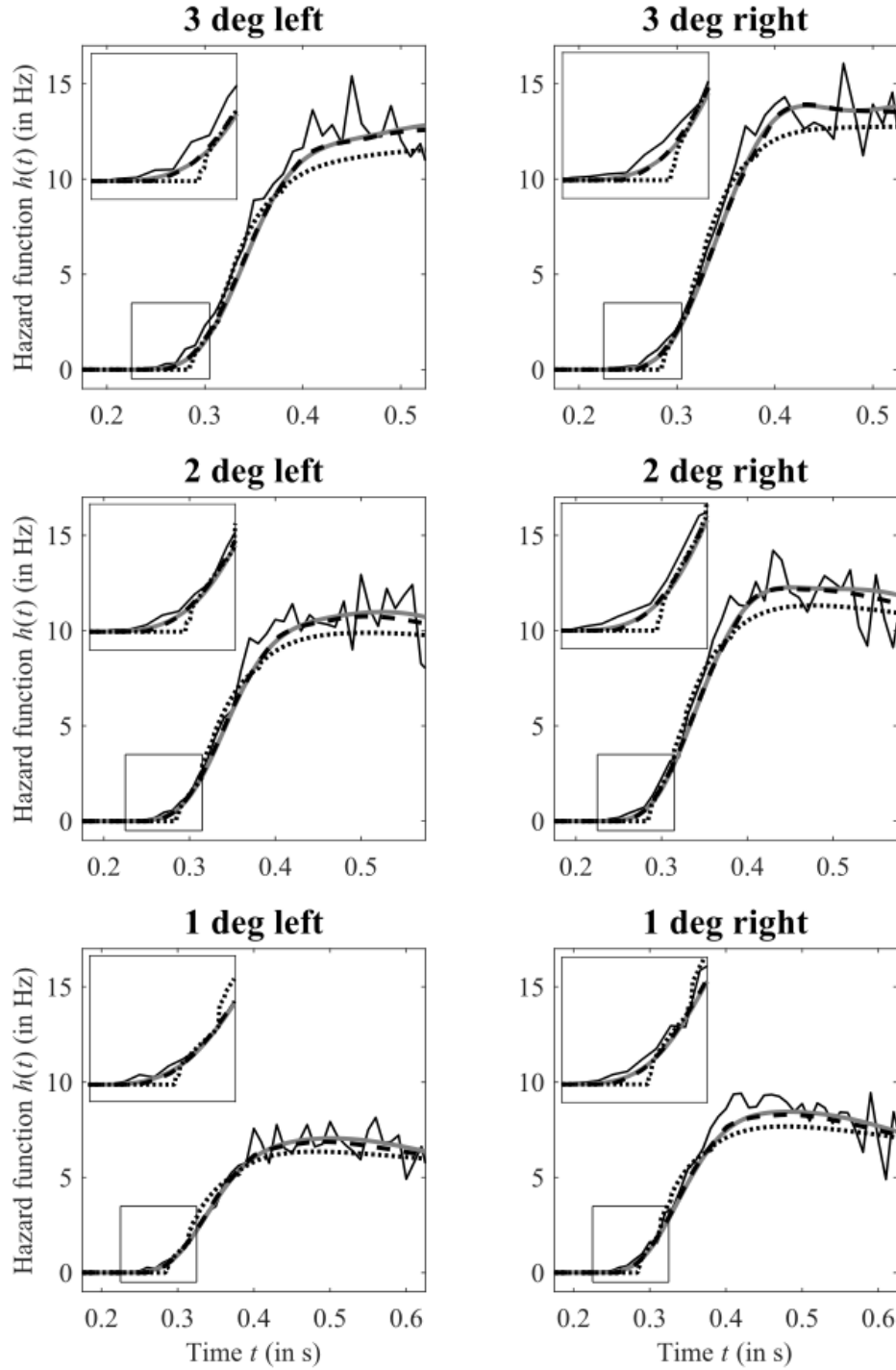


Figure 7. Empirical (thin black line) and predicted hazard rates (bold lines) of response times. Predicted hazard rates are obtained from the Poisson random walk with time inhomogeneous (solid gray line) or constant processing rates with variable T_0 (dashed black line) or constant T_0 (dotted black line). The time inhomogeneous model particularly gives a more accurate description of the smooth increase at the leading edge of all RT distributions than the model with constant processing rates and constant T_0 . Predictions of the constant processing rate model with variable T_0 are very similar to those of the time inhomogeneous model. Model parameters of the time inhomogeneous model are summarized in Table 1.

Fit of the Poisson counter model to pure accuracy data

We now turn to the question whether the processing rates and therefore the concept of a categorization “item x belongs to category i ” that the RT model is based on is the same as in TVA. We will demonstrate that this is indeed the case by applying TVA for pure accuracy tasks with confusable stimuli; that is, a Poisson counter model to the stimuli used in Experiment 1 in a pure accuracy task with backward masking. First, we provide the results of an independent fit, replicating and extending results reported by Kyllingsbæk et al. (2012). Then, we compare the results obtained in separate fits of the Poisson random walk (RT) and the Poisson counter model (pure accuracy) to a combined fit model of both datasets.

The pure accuracy data was modelled with the Poisson counter model described in Kyllingsbæk et al. (2012). This version of a Poisson counter model does not entail any response criteria (Pike, 1973). Rather, a set of response rules maps the states of the assumed Poisson counters to observable behavior. These response rules are based on the relative evidence of counters and since there are only two response categories, the Poisson counter model could as well be formulated as a random walk model⁴. Although this representation is certainly instructive to inspect the many commonalities of the two models, the random walk representation is limited to the special case of two perceptual categories, so we used the more general formulation with Poisson counters.

We slightly adapted the Poisson counter model as described in Kyllingsbæk et al. (2012) to match the assumptions we made for the Poisson random walk—that is, trial-by-trial variability in processing rates and time inhomogeneous processing. We briefly summarize the main assumptions of the Poisson counter model of categorization in a pure accuracy task.

⁴ In the special case of only two response alternatives, identical model predictions can be obtained from the Poisson counter and Poisson random walk model. For example, an equal amount of counts in the counter model corresponds to being in the neural state in the random walk. We prefer to use the term Poisson counter model, however, because this equivalence only holds for two response alternatives and only for the unrestricted process. This does not imply that we assume that participants shifted their strategy between conditions.

Then, we outline the adaptations made to obtain a version of the model that is comparable to the Poisson random walk.

The main assumptions underlying the Poisson counter model are the same as those of the Poisson random walk model of RT: Categorization of mutually confusable stimuli is conceived as a sequential sampling process. It is assumed that the reported perceptual category is based on the amount of evidence obtained for each perceptual category i in form of tentative categorizations (instead of the first categorization that finishes processing). Tentative categorizations are made as long as stimulus information is available—that is, until mask onset. The observer is further assumed to carefully evaluate the evidence for each perceptual category and finally report a perceptual category according to the following rules: First, if one category has accumulated more evidence than any other, this category is reported. Second, if two or more categories have the same amount of evidence, but more than all others, the observer randomly (i.e., uniformly) selects one of those categories (e.g., Ratcliff, 1988). Finally, if no evidence has been accrued for any perceptual category, the observer refrains from responding or, with a certain probability, responds by guessing blindly among all perceptual categories.

If the processing rate $v(x, i)$ is constant over time, the probability that n categorizations are made for perceptual category i until time t follows a Poisson distribution:

$$P[n \mid v(x, i)t] = \frac{[v(x, i)]^n t^n}{n!} \exp[-v(x, i)t] \quad (8)$$

The model is easily generalized to include trial-by-trial variation in processing rates. If we assume the same rate variation as in the Poisson random walk; that is, processing rates are gamma-distributed with parameters $\theta(x, i)$ (shape) and $1/\xi$ (scale), the probability for n categorizations—given by Equation (8) for the model with constant rates across trials—follows a negative binomial distribution instead:

$$P[n \mid \theta(x, i), \xi, t] = \frac{\Gamma[n + \theta(x, i)]}{n! \Gamma[\theta(x, i)]} \frac{t^n \xi^{\theta(x, i)}}{(t + \xi)^{n + \theta(x, i)}}$$

As is the case with Gaussian increments (Ratcliff, 1978), this assumption prevents stimulus discriminability to become infinite with unlimited exposure duration. Rather, it depends on the shape and scale parameter of the distribution of processing rates:

$$d' = \frac{E(X_A - X_B)}{\sqrt{\text{Var}(X_A - X_B)}} = \frac{\theta_A - \theta_B}{\sqrt{(\theta_A + \theta_B)\xi}} \times \frac{t}{\sqrt{t^2 + \xi t}}$$

In the limit ($t \rightarrow \infty$), the second term on the right-hand side will approach unity. Thus, in contrast to a model with constant processing rates, the predicted frequency of errors will not generally go down to zero with increasing exposure duration. As shown above, the Poisson distribution can be generalized to account for time-inhomogeneous processing (Smith & Van Zandt, 2000). The model with both trial-by-trial variability in processing rates and time inhomogeneous processing is described in Appendix A.

For the fit of the Poisson counter model to pure accuracy data, we specified the processing rates as in the Poisson random walk model for RT data: $\bar{v}(x, i) = \eta_\theta(x, i) \beta_\theta(i)/\xi$. As in the RT model, we restricted the perceptual evidence for correct ($\eta_\theta(x, C)$) and incorrect categorizations ($\eta_\theta(x, I)$) to be the same for both perceptual categories. Similarly, the higher of the two bias parameters was set to unity and the lower of the two was a free parameter (determined on an individual basis)

The Poisson counting process was modelled as a temporally inhomogeneous process. In contrast to the RT model we included contributions of transient channels in the model of pure accuracy (e.g., Christensen, Markussen, Bundesen, & Kyllingsbæk, 2018; Petersen & Andersen, 2012). Perceptual evidence obtained from transient encoding channels was assumed to be the same for all categorizations. Finally, the response rules required two

guessing parameters to be estimated from data, that is, the probabilities p_A and p_B ($p_A + p_B \leq 1$) to guess any of the two categories when no evidence was obtained at all. The probability of not reporting a perceptual category (trial rejection) is therefore $1 - p_A - p_B$. Altogether, the model had 13 parameters to explain data (accuracy of report) in 60 conditions (exposure duration \times stimulus orientations) and performance in catch trials.

The fit of the Poisson random walk model to pure accuracy data was again excellent. The model explained on average 96.7% (95.2 to 97.8%) of variation. The predictions of the model closely matched the empirical data in all conditions, differences between model and data were minimal and model predictions always lay within the 95% confidence interval bounds (as determined by bootstrap). This was also true for the predicted and observed frequency of no report (unforced choice). A critical feature of the Poisson counter model for categorization of mutually confusable stimuli is its ability to predict the non-monotonic probability of selecting an incorrect categorization. While the probability of a correct categorization increases with increasing exposure duration, the probability of an incorrect categorization first increases, and then decreases. This pattern is more pronounced with higher stimulus confusability (Figure 8).

Table 2.

Group average ($\pm SD$) of parameter estimates for the Poisson counter model of accuracy of report in the pure accuracy task.

Parameter			Group average ($\pm SD$)	
			Correct categorization	Incorrect categorization
$\bar{v}(x, i)$ (in Hz)	Stimulus orientation	3 deg left	71.1 (± 17.2)	2.9 (± 4.7)
		2 deg left	63.3 (± 14.5)	6.5 (± 6.3)
		1 deg left	43.8 (± 10.1)	12.2 (± 5.4)
		3 deg right	76.5 (± 20.8)	2.5 (± 4.2)
		2 deg right	68.7 (± 20.4)	6.1 (± 6.2)
		1 deg right	47.7 (± 14.4)	11.1 (± 4.7)
$\eta_{\theta}(x, i)$	Stimulus orientation	3 deg	23.7 (± 27.6)	0.6 (± 1.2)
		2 deg	22.3 (± 27.8)	3.5 (± 5.8)
		1 deg	14.6 (± 17.4)	4.5 (± 6.2)
$\beta_{\theta}(i)$	Perceptual category	“left”	0.89 (± 0.11)	
		“right”	0.95 (± 0.08)	
$1/\xi$			7.43 (± 4.12)	
μ_S			441 (± 298)	
μ_T			429 (± 261)	
$\eta_{\theta,T}$			0.185 (± 0.254)	
p_i	Guessing	“left”	.164 ($\pm .198$)	
		“right”	.211 ($\pm .221$)	

Note—Bar parameters $\bar{v}(x, i)$ denote mean processing rates (mean of gamma distribution with scale $1/\xi$). Mean processing rates were not free parameters but obtained from perceptual evidence $\eta_{\theta}(x, i)$ and bias $\beta_{\theta}(i)$ according to Equation (7). One of the bias parameter was fixed at unity (the highest, individually for each participant). Parameters μ_S and μ_T specify the buildup of Poisson processing rates in sustained and transient channels, respectively (time inhomogeneous processing). The mean processing rate of transient channels is $\bar{v}_T = \eta_{\theta,T} \beta_{\theta}(i) / \xi$, irrespective of x . If no evidence was obtained, p_i is the estimated probability of guessing perceptual category i . The probability of not responding is $1 - \sum p_i$. The standard deviation of parameters (in parentheses) was obtained across participants.

The processing rates for both correct categorizations ($F_{2,18} = 31.46, p < .001$) and incorrect categorizations ($F_{2,18} = 16.11, p < .001$) were significantly different across confusability conditions. In line with expectation and the results obtained in speeded RT, perceptual evidence for the correct categorization decreased with increasing confusability, whereas perceptual evidence for the incorrect categorization increased with confusability. The

processing rates for correct categorizations in low confusability were significantly higher than those in intermediate ($t_{18} = 2.14, p = .047$) and high confusability conditions ($t_{18} = 7.69, p < .001$). These differences were 7.8 Hz (95%-CI: 0.7 to 14.9 Hz) and 28.1 Hz (20.9 to 35.2 Hz), respectively. Processing rates for intermediate confusability were also significantly higher (20.3 Hz, 13.1 to 27.4 Hz) than for the high confusability conditions ($t_{18} = 5.55, p < .001$). Conversely, processing rates for incorrect categorizations of highly confusable targets were significantly higher than for intermediate confusability ($t_{18} = 3.40, p = .003$) and low confusability targets ($t_{18} = 5.64, p < .001$). These differences were 5.4 Hz (2.3 to 8.5 Hz) and 9.0 Hz (5.9 to 12.1 Hz), respectively. The rate for incorrect categorizations in intermediate confusability was significantly higher ($t_{18} = 2.24, p = .038$) than in low confusability targets (3.6 Hz, 0.5 to 6.7 Hz).

There was a slightly higher bias for the categorization “item x is oriented rightwards” ($\beta_{\theta}(\text{left}) = 0.89, \beta_{\theta}(\text{right}) = 0.95$). However, this difference in perceptual bias, as assessed by the processing rates for “leftwards” and “rightwards” categorizations, was not significant ($t_9 = 1.26, p = .238$). The average bias in processing rates for “rightwards” categorizations across all confusability levels and correctness of the categorization was 2.8 Hz (−1.7 to 7.3 Hz). The effect of transient encoding channels was crucial to explain performance at short and medium exposure durations, especially to explain the increasing probability to commit an error. Activity of the transient encoding channels ($\mu_T = 429$) peaked quickly (at about 10 ms⁵). Although the estimated perceptual evidence obtained from transient channels was low compared to the sustained channels, their contribution was substantial, particularly their contribution to erroneous categorizations

⁵ Obtained from the mode of a gamma distribution with rate μ_T and shape $n_T = 5$ (fixed)

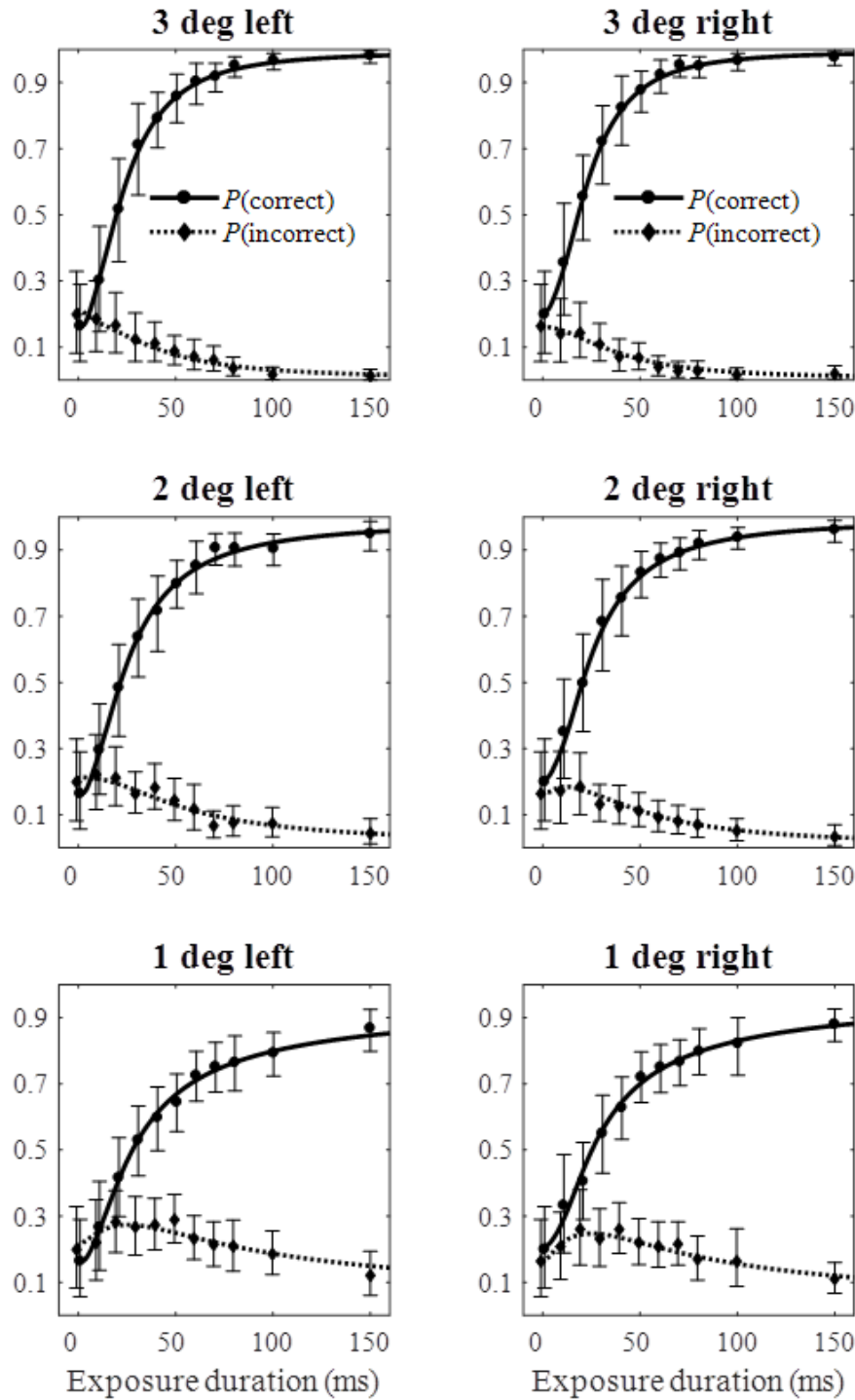


Figure 8. Fit of the Poisson counter model for pure accuracy. Observed percentages of correct and incorrect responses dependent of exposure duration are represented by circles and diamonds, respectively. Model predictions for the probability of a correct report are displayed as unmarked points connected with solid lines; predictions for incorrect reports are displayed as unmarked points connected with dashed lines. Model parameter estimates are summarized in Table 2. Error bars denote 95% confidence intervals (determined by bootstrap).

At long exposure durations (100 ms and 150 ms), performance was nearly perfect in the two conditions with less confusable stimuli, whereas selecting the incorrect perceptual category occurred frequently at these exposure durations in the high confusability condition. Despite the instruction not to guess and in contrast to previous application of the Poisson counter model (Christensen et al., 2018; Kyllingsbæk et al. 2012), participants reported either categorization in almost 40% of catch trials, on average. This behavior transfers directly into the model by means of the guessing parameters p_{left} (.164) and p_{right} (.211). In other words, these parameters were almost exclusively determined by the guessing rates observed in the catch trials. According to the Poisson counter model fit, the probability to report a category in absence of information about the target was significantly higher than zero ($t_9 = 3.07$, $p = .013$). However, the probabilities p_{left} and p_{right} for guessing in absence of any information about the target were not significantly different ($t_9 = 0.92$, $p = .381$).

The individual fits suggest that participants adopted quite different response strategies regarding trial rejection. This can be seen in the data with large confidence intervals on the left ($ED = 0$) and the standard deviation for the parameter representing the probability to guess among categories. We suspect that this variation reflects decision strategies adopted by the participants in case they had not seen the target (i.e., guessing) and for trials in which they had seen the target but were not certain enough to give a response—that is, rejecting the trial despite having obtained at least some evidence in favor of one or the other category.

The obtained estimates suggest that processing rates were very similar to the ones obtained from the data of the speeded response time task. The temporal profile of encoding, however, was different between the two tasks. In speeded responses, input from sustained encoding filters responded rather slowly (see above), whereas in pure accuracy, the input from sustained encoding filters increases faster. Guessing varied greatly across participants, the probability of no report was between 95% and almost zero, with a slight bias to guess the

perceptual category “oriented rightwards” in absence of evidence, in line with the overall bias for that perceptual category observed in Experiment 1.

A combined model of categorization

Although the two models explain performance in very different tasks, the underlying assumptions are very similar: evidence is collected in form of tentative categorizations and the amount of evidence for each perceptual category is stored in a Poisson counter. Also, the response rules that map these latent processes onto observable responses are based on the difference of two counters in either case. This similarity means that the two models share central parameters, namely those describing the buildup of evidence: the Poisson processing rates and the model parameters they are derived from. As can be seen in Tables 1 and 2, the models share the perceptual evidence parameters ($\eta_\theta(x, i)$), bias parameters ($\beta_\theta(i)$), the scale parameter of processing rates ($1/\xi$), and the parameter of the response function of sustained encoding channels (μ_S). Before presenting the results of a combined model for pure accuracy and speeded RT in which these parameters were constrained to be equal, we investigate more closely how well the independent fits match. For this, we re-fitted both models to an average observer to eliminate inter-subject variation.

The mean processing rates show the same pattern across difficulty, that is, with increasing confusability processing rates for incorrect categorization increased and processing rates for correct categorizations decreased. The processing rates of correct categorizations are almost identical in both tasks, ranging from about 45 to 75 Hz. In contrast, processing rates of incorrect categorizations were rather low in the pure accuracy task and consistently lower than those obtained from RT data. This is due to the transient processing in the Poisson counter model which contributes significantly to the accumulation of evidence in favor of the incorrect perceptual category. To make the models comparable, we eliminated transient processing from the pure accuracy model and instead modified the response rule so that more

than one tentative categorization may be necessary to conclusively select and report any perceptual category. The number of parameters remains the same and the fit ($R^2 = 0.995$) to the average observer is very similar to the fit presented in the previous section ($R^2 = 0.995$).

Table 3.

Group parameter estimates determining processing rates in the Poisson counter model for the pure accuracy task and Poisson random walk model for speeded responses.

			Pure accuracy categorization		Speeded RT categorization	
Parameter			Correct	Incorrect	Correct	Incorrect
$\bar{v}(x, i)$ (in Hz)	Stimulus orientation	3 deg left	54.0	4.9	65.9	14.4
		2 deg left	51.0	6.2	59.7	15.5
		1 deg left	41.4	12.1	47.4	20.4
		3 deg right	57.2	4.6	72.2	13.1
		2 deg right	54.0	5.9	65.5	14.1
		1 deg right	43.8	11.4	52.0	18.6
$\eta_\theta(x, i)$	Stimulus orientation	3 deg	8.0	0.7	18.7	3.7
		2 deg	7.6	0.9	16.9	4.0
		1 deg	6.2	1.7	13.5	5.3
$\beta_\theta(i)$	Perceptual category	“left”	0.94		0.91	
		“right”	(Fixed at 1)		(Fixed at 1)	
$1/\xi$			7.13		3.86	
μ_s			753		50.0	

Note—Mean processing rates $\bar{v}(x, i) = \eta_\theta(x, i) \beta_\theta(i)/\xi$ represent mean of a gamma distribution with scale $1/\xi$. Parameters μ_s specify the buildup of Poisson processing rates (time inhomogeneous processing). On group level, $\beta_\theta(\text{right})$ is fixed at unity ($\beta_\theta(\text{right}) > \beta_\theta(\text{left})$). Altogether, the two models share nine free parameters: six perceptual evidence parameters ($\eta_\theta(x, i)$), a perceptual bias parameter ($\beta_\theta(\text{left})$), the scale parameter of the processing rate distributions ($1/\xi$), and the sustained encoding filter parameter (μ_s).

The RT model fitted to the average observer was identical to the one fitted to individual observers. Even so, eliminating inter-subject differences further increased the goodness-of-fit to an almost perfect account of the averaged data ($R^2 = 0.999$). The relevant parameters for the comparison of both models are summarized in Table 3.

We first tested a model in which all common parameters (i.e., all $\eta_\theta(x, i)$, $\beta_\theta(\text{left})$, ξ , and μ_s) were restricted to be equal. This worsened the model fit to speeded RT ($R^2 = 0.968$) whereas the model fit to pure accuracy data was about the same ($R^2 = 0.995$). Altogether,

model fitted significantly worse than the separate fits ($\chi^2 = 115.3$; $df = 9$; $p < .001$). It turned out that the main reason for this lack of fit was the difference in sustained encoding channels (μ_S) between the models: the input from encoding channels increased much more quickly in the pure accuracy task ($\mu_S = 753$) than in the speeded response task ($\mu_S = 50$). When this difference was accounted for, the model fitted about equally well as the separate models (RT: $R^2 = 0.996$; Accuracy: $R^2 = 0.995$) and the lack of fit was no longer significant ($\chi^2 = 11.69$; $df = 8$; $p = .166$). No other parameter had the same impact and the difference in μ_S was significant even if μ_S was the only parameter being restricted ($\chi^2 = 72.2$; $df = 1$; $p < .001$). Taken together, we conclude that the perceptual processing parameters were approximately equal in both tasks except that the encoding filters reacted more quickly after stimulus onset in the pure accuracy task than in the speeded response time task.

The pronounced temporal difference in encoding was unexpected, so additional analyses were performed to shed light on its causes. For one of these, we split the data into three datasets collected in the three sessions for each participant. We then re-fitted both Poisson models to pure accuracy data and speeded RTs with a separate μ_S for each session to obtain an approximate temporal profile of this parameter. It turned out that there was a pronounced increase in the pure accuracy tasks (Session 1: $\mu_S = 355$; Session 2: $\mu_S = 873$; Session 3: $\mu_S = 929$), but not in the speeded RT task (Session 1: $\mu_S = 39.7$; Session 2: $\mu_S = 51.2$; Session 3: $\mu_S = 55.6$). It seems that the participants adapted to the temporal demands of briefly presented targets by tuning their encoding channels to its onset, and that this learning was a main reason for the difference in the temporal profile of encoding channels between the pure accuracy task with briefly presented targets and a speeded RT task with response terminated targets.

Discussion

Applications of the theory of visual attention (TVA) over the last decades has focused on accuracy of report in partial and whole report paradigms (Bundesen & Habekost, 2008). TVA is formalized as a simple race model according to which the first K categorizations for x are stored in VSTM. With only a single stimulus, the first categorization that finishes processing and is encoded in VSTM is the one that is selected (Bundesen & Harms, 1999). With this simple framework, it is possible to explain all aspects of the accuracy of report; its parameters and their properties have been thoroughly studied (Habekost et al., 2014).

It is possible to derive predictions for RT from within this framework as well (Bundesen, 1990): Predictions for mean RT have been considered, for example, for the spatial cueing task (Posner et al., 1980) and feature search (Treisman & Gelade, 1980). In principle, the TVA framework could be used to explain RT distributions as well; however, the simple race model structure of TVA is not sufficient for this purpose. In this article, we extend TVA to a sequential sampling model of RT and attention which contains TVA as a special case (i.e., when the evidence threshold k is unity). Common sense as well as the data obtained in Experiment 1 speak against a simple race model to account for RT distributions in speeded categorization of confusable stimuli. We have reasoned that stimulus confusability leads to repeated categorization of stimulus x —a notion that was confirmed by the results of Experiment 1.

The proposed Poisson random walk model describes all aspects of performance in a speeded RT task with two response alternatives and parameter estimates are comparable to those obtained in a pure accuracy categorization task. With standard assumptions in RT modeling, that is, trial-by-trial variation in processing rates and starting point, the Poisson random walk model could predict both correct and error RT distributions. A combined fit of the RT model and a Poisson counter model for pure accuracy revealed that model parameters describing the processing rates of categorizations for x in speeded categorization were highly

similar to those obtained in a non-speeded, non-forced categorization task. We conclude that the Poisson random walk for response times describes speeded categorizations with the same underlying cognitive processes as TVA—repeated categorization appears to be a suitable way of generalizing TVA to situations in which stimulus categorizations are mutually confusable.

Both the Poisson counter model and the Poisson random walk model resolve the problem of decision noise with the assumption that a categorization is selected based on the accumulated evidence in favor of that perceptual category, rather than on the first categorization for x that finishes processing. The Poisson random walk model is a generalization of TVA in the sense that if more than one perceptual category is active, repeated sampling of stimulus evidence can overcome decision noise. It has been noted that TVA in most cases describes a noise-free representation of the visual stimuli—or, at least that encoding noise is negligible (Logan, 2004). Assuming no decision noise, errors occur only as misses if encoding is not finished before processing of the target is interrupted by the mask. The proposed Poisson random walk model offers two sources of noise. The first is decision noise which arises when more than one template is active in any trial: this is caused by confusability of the stimuli so that it is difficult to determine whether “item x belongs to category i ” or “item x belongs to category j ” is the correct categorization. This leads to within trial noise described by the random increments and decrements of the random walk. The second source of noise is perceptual noise. This type of noise is dependent on the quality of the presented stimuli. Noisy or low-contrast stimuli would lead to lower processing rates, slowing down processing and causing slower responses. In the model, decision noise (confusability) and perceptual noise (visibility) can be accounted for independently. However, perceptual noise is often confounded with decision noise because degraded stimuli are often less distinguishable. Therefore, we used high-contrast, mutually confusable stimuli in Experiment 1 to minimize perceptual noise.

The performance in catch trials was highly different between pure accuracy and speeded RT. As stated above, a category was frequently reported in the pure accuracy task, whereas all participants could refrain from responding in speeded RT. Catch trials for speeded RT are common in experiments involving a detection task (simple response), but rather uncommon in discrimination tasks. One reason to include catch trials in the speeded RT condition was to keep the experimental design balanced across speeded RT and pure accuracy. Another, more implicit, reason to include catch trials is to discourage fast guesses. Due to the task demands in both conditions, the catch trials were rather different, so that the difference in observable behavior is not too surprising. From a modeling point of view, one can assume that processing rates for categorizing a target that has not been shown is zero, or at least very low. This is what we assumed when modeling the pure accuracy task. Under the same assumption it is very unlikely in the speeded RT task that a response threshold is reached during the duration of the catch trial. When no threshold is reached in time, no response is produced—which is what we observed in the speeded RT condition of Experiment 1.

In contrast to Poisson counter models (Kyllingsbæk et al., 2012; Pike, 1973; Townsend & Ashby, 1983), the idea of a Poisson random walk model has received little attention in the past. Notable exceptions are the spiking sequential probability ratio test (SPRT) model of Zhang and Bogacz (2010) and the exemplar-based random walk (EBRW) of Nosofsky and Palmeri (1997a/b; Palmeri, 1997). The EBRW was proposed to model response time data in an object classification task. The rates of the underlying Poisson processes were assumed to reflect the distance between an object and an object class. In the initial description of their model, the expected number of steps in the random walk was considered which gives rise to mean RT predictions (Nosofsky & Palmeri, 1997a; Palmeri, 1997; Zhang & Bogacz, 2010); later, the model was extended to account for RT distributions (Nosofsky & Palmeri, 1997b; see also Logan, 2002; Nosofsky, Little, Donkin, & Fific, 2011). Our model aims at providing

a full account of response times in speeded categorization of confusable stimuli—that is, predictions for the distribution of correct RTs as well as for error RT distributions.

Although the fit of the model to the data is excellent, there are limitations that might speak against the model. One of these is that in our attempt to use the same stimuli for both a response time and a pure accuracy task, the range of task difficulty that could be used was rather restricted. Usually, response time models are applied to data containing conditions with lower response accuracy than we observed in Experiment 1. This was not possible here, the average response accuracy in the most difficult condition (1 deg Gabor patches) was about 87%, so it can be argued that the model might fail to predict response time distributions in experiments with a wider range of task difficulty. Secondly, we did not collect data to assess effects of speed-accuracy instructions on response time distributions directly—one of the main advantages of employing explicit RT models such as the diffusion model.

To address these concerns, we fitted the model to data from an RT study (Ratcliff, Thapar, & McKoon, 2001) to which the diffusion model had previously been fitted. The data of this study already served as a benchmark for a selection of RT models and some models (including the Poisson counter model) had difficulties explaining the various effects of speed-accuracy instructions on the pattern of RT distributions (Ratcliff & Smith, 2004). Details on this comparison can be found in Appendix D. From this fit, it is evident that the Poisson random walk model can explain RT distribution in conditions with considerably lower response accuracies than those we observed in Experiment 1. Response accuracy in the two most difficult conditions (speeded instruction: 69% and 58%, accuracy instruction: 74% and 60% for the younger participants) is far lower than the most difficult condition in our experiment. However, there is no systematic mismatch between model and data in these conditions. Secondly, the Poisson model is also well capable of accounting for a speed-accuracy trade-off. Different evidence thresholds k accounted for differences in RT distributions under speed and accuracy instructions. In sum, we conclude that the Poisson

random walk model is as capable as the diffusion model in its ability to account for response time distributions of correct and error responses under varying speed/accuracy instructions.

Visual processing capacity

Over the past decades, response times have become a major dependent variable in cognitive psychology. Response times can be analyzed on basis of mean RT in a purely statistical manner using statistical tests to compare response time across a set of experimental conditions. Way fewer studies have used complete response time distributions for analyses. Of these, the two main directions are explicit RT modelling accounts (e.g., Brown & Heathcote, 2005, 2008; Link & Heath, 1975; Ratcliff, 1978; 1981; Smith & Ratcliff, 2009; Usher & McClelland, 2001) and non-parametric analyses of survivor, or, hazard functions (e.g., Koch et al, 2013; Miller, 1982; Townsend & Ashby, 1983; Townsend & Nozawa, 1995; Townsend & Wenger, 2004). These approaches were developed quite independently of each other; the most prominent examples of each class—the diffusion model and the systems factorial technology (SFT, Townsend, 1992)—have not been directly related to each other (for the LBA model, see Eidels, Donkins, Brown & Heathcote, 2010). The diffusion model is void regarding processing capacity, whereas SFT can be used to compare cognitive architectures under different assumptions, such as serial vs. parallel (e.g., Yang, Little, Hsu, 2014) or facilitatory vs. inhibitory parallel processing (Eidels, Houpt, Altieri, Pei, & Townsend, 2011; Townsend & Wenger, 2004). However, it does not describe the exact cognitive process that produced the observed pattern of results (Townsend & Nozawa, 1995; but see Eidels et al., 2011). The proposed Poisson random walk response time model is suitable to bridge this theoretical gap. It accounts for processing capacity and the model can be readily applied to speeded responses to a varying number of stimuli. So far, only single stimulus recognition was considered. By including the TVA assumption on attention weights w_x (see Equation (2)), the Poisson random walk can be used to model speeded responses to a varying number of

targets and distractors including the effect of multiple targets and/or distractors on visual processing capacity.

Varying the number of targets and distractors is also central to classic paradigms to which TVA applies: in partial report paradigms, the number of distractors presented along targets is systematically varied. With this variation it is possible to investigate whether processing capacity is limited, meaning that when additional targets are added to the display the processing of these targets slows down because processing capacity is distributed across more processes. On the other hand, processing capacity could be unlimited, meaning that additional targets activate different processing resources so that the overall processing capacity increases with the number of targets in the display. The question of processing architecture (i.e., serial or parallel; self-terminating or exhaustive) could be addressed by the workload capacity index (Little, Eidels, Fific, & Wang, 2015). Future experimental work is needed to empirically test the effect of multiple target stimuli (e.g., redundant signals, Schwarz, 1989) or multiple distractors (e.g., visual search, Banks & Prinzmetal, 1976) on visual processing capacity obtained from RT distributions in speeded RT tasks.

Speeded categorization with multiple stimuli

In Experiment 1, we have tested the PRW model with data from single stimulus recognition. The rate equation inherited from TVA (Equation (2)) makes the model applicable to a much wider set of experimental paradigms. One of these is the presentation of distracting items along with a target. In single stimulus recognition, the weight of the target in the rate equation can be set to unity ($w = 1$) without loss of generality. By Equation (2), visual processing capacity is then determined by the processing rates of the target. When distractors are presented alongside the target, this may no longer be the case. The presence of the distractors will lower the processing speed of the target, if one assumes a fixed processing capacity and unless filtering is perfect (in which case distractors get a weight of zero).

Filtering and the selection of objects from multielement displays are primary aspects of TVA, however, describing this selection process in terms of RT leads to rather complex models in which processing of each item is described by a random walk (Wolfe, 2007). Fortunately, in some situations, it may be appropriate to apply some simplifying assumptions. For example, in a categorization task in which the distractors are not too similar to the target, it may be appropriate to assume that observed RTs are based on the outcome of processing of the target, but not distractors. What are the predictions of the Poisson random walk model in such a categorization task? Under the assumption of fixed capacity, adding distractors to the stimulus display will decrease the processing speed of the target since its relative attention weight (i.e., the fraction of weights w in Equation (2)) will decrease with the number of distractors in the display. With fixed capacity, the model predicts slower responses when the number of distractors increases. Interestingly, predicted response accuracy is completely unaffected by the number of distractors. This strong prediction matches results reported by previous studies of target discrimination with distractors (Bacon & Egeth, 1994; Theeuwes, 1990). The prediction also gains support from results of visual search with mere target detection: Depending on the search task, response times frequently increase with set size, whereas accuracy often stays relatively constant⁶ (e.g., Duncan & Humphreys, 1989, 1992; Treisman & Gelade, 1980; Wenger & Townsend, 2006; Wolfe, Palmer, & Horowitz, 2010).

Predicting set size effects on RT—but not accuracy—is due to the interplay between pigeonholing and filtering assumed in the processing rates for the PRW. Changing the attention weights of an item (filtering) will not change the relative size of the processing rates for categorizing the target correctly or incorrectly (pigeonholing). Instead, if the weight of an object changes, this changes the overall processing rate of an item (i.e., the sum of processing rates for each response category) but leaves their relative strength unchanged. This property is

⁶ With constant, we refer to the observation that error rates barely increase and often do not exhibit a monotonic relationship with set-size. Additional assumptions, such as terminating the search prematurely upon reaching a deadline (e.g., Bundesen, 1990) may explain small increases in miss rates, if present.

inherited from TVA; according to TVA, pigeonholing and filtering work independently from each other—that is, changing attention weights does not affect pigeonholing and changing bias does not affect filtering (Bundesen, 1990). In contrast to TVA (which has a response criterion of unity), RT predictions additionally depend on the response criteria, so that the slowdown due to distractors becomes more prominent with higher response thresholds. With a given response threshold, however, response accuracy remains independent of the number of distractors. This contrasts with, for example, the role of drift rates in the diffusion model of response times (Ratcliff, 1978). Changing the drift rate changes predictions of both response time and accuracy (i.e., if responses get slower, they also will be less accurate).

To illustrate the predictions of the PRW model, consider an experiment like Experiment 1 (orientation discrimination), but with n_D distractors added to the display (e.g., $n_D \in \{0, 1, 3\}$). We assume a fixed processing capacity of $C = 80$ Hz, no bias ($\beta = 1$), an attention weight of unity for the target ($w_T = 1$) and a lower attention weight for distractors ($w_D = 0.2$). Further, assume four confusability conditions ($j = 1, 2, 3, 4$) with increasing confusability in stimuli x_j , with $\eta(x_j, 1) = [60, 52, 48, 44]$ for Category 1, defined as the correct report category. With response criteria set to $k_a = k_b = 3$, response accuracy is about 96%, 87%, 77%, and 65%, respectively, irrespective of the number of distractors. The number of distractors has, however, an effect on the predicted response time distributions (Figure 9). Predicted response times are increasing (i.e., responses get slower) with the number of distractors in the display.

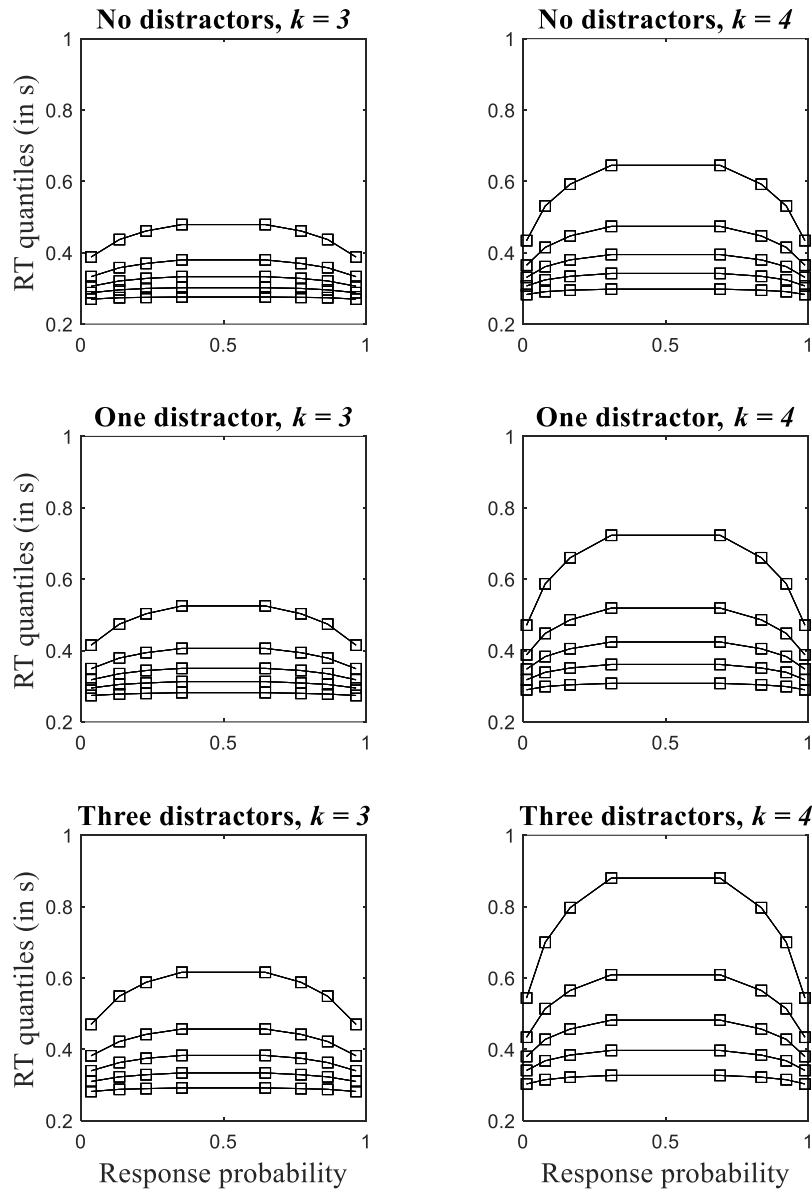


Figure 9. Model predictions for a single target presented without distractors (top row), with a single distractor (middle row) or with three distractors (bottom row). The panels in the left column show a less cautious, those in the right column a more cautious response strategy. With a fixed processing capacity C , the number of distractors has a clear effect on RT predictions (ordinate, slowdown from top to bottom,), which is more pronounced with higher response thresholds. By contrast, the predicted accuracy (abscissa) does not change with the number of distractors.

With a more cautious response strategy ($k_a = k_b = 4$), accuracy increases and responses get slower, compared to the example above with lower criteria. Again, the number of distractors has no effect on predicted accuracy (99%, 92%, 84%, and 61%, respectively),

whereas the predicted slowdown of responses is more even pronounced than with lower criteria. This is just one example of strong and testable predictions of our TVA-based RT model. Clearly, this simple model will fail, if, for example, the distractors themselves are so similar to the target stimulus that the probability of incorrectly responding to one of the distractors can no longer be neglected. Further investigations are needed to test whether these PRW predictions for responses to targets in multi-element displays are supported by data.

The PRW as a cognitive model of speeded categorization

Both visual processing capacity C and the processing rates have a clear cognitive interpretation (Table 4). They express processing speed in terms of categorizations per second (Hz). In physical terms, they represent the power of a system; the amount of work that is done per time. Related to these is the integrated event rate (see Figure 4), the expected number of categorizations that have occurred up to time t . The integrated event rate corresponds to the work done in a time interval. Both these concepts were generalized to time dependent processing rates in the Poisson random walk.

Table 4.

Description of parameters of the Poisson random walk

Symbol	Name	Description
$v(x, i)$	Processing rate	Mental speed of processing the categorization “item x belongs to category i ”. These are composed of perceptual evidence and bias: $v(x, i) = \eta(x, i)\beta_i$.
$\eta(x, i)$	Perceptual evidence	Strength of perceptual evidence for the categorization “item x belongs to category i ”. The more similar an item is to category i , the higher the perceptual evidence.
β_i	Category bias	Category bias for category i , independent of stimulus x .
C	Visual processing capacity	Overall mental speed of processing. Processing capacity is the sum of all processing rates.
$\Lambda(t)$	Integrated event rate	Number of expected categorizations at time t . Sum of mental work done up to time t . If processing rates are constant, $\Lambda(t) = Ct$.
k_i	Response criterion/ Evidence threshold	Number of excess categorizations needed in favor of perceptual category i to be selected.
$\eta_\theta(x, i)$	Perceptual evidence [trial-by-trial variation]	Strength of perceptual evidence with gamma distributed processing rates (shape parameter).
$\beta_\theta(i)$	Category bias [trial-by-trial variation]	Category bias with gamma-distributed processing rates (shape parameter).
$1/\xi$	Distribution parameter for $v(x, i)$	Scale parameter of the gamma distribution of processing rates (trial-by-trial variation). The higher the scale parameter, the higher the variance compared to the mean.
z_i	Starting point distribution [trial-by-trial variation]	Describes the variation of starting point. If all z_i are equal, the starting point varies, but is unbiased. The lower the value, the higher the variation. the greater the difference between z_i , the higher the initial bias towards categories i with high values.
μ_s/n_s	Sustained encoding filter response	Describe the temporal characteristics of sustained encoding channels. Higher parameters represent slower responses. Filter amplitude is given by $v(x, i)$.
T_0	Residual time/ Non-decision time	Latency of all processes otherwise not covered by the model.

The evidence threshold k_i represents the number of excess tentative categorizations that are needed to conclusively select perceptual category i . If evidence thresholds are high, responses are slower and more accurate. Evidence thresholds are assumed to be set independently of perceptual processes⁷, so which case it is possible to distinguish between response bias, or, response caution and perceptual processing (sensitivity). Based on the data from Experiment 1 and the comparison with the diffusion model (Appendix D), the number of excess tentative categorizations is around $k = 4$ when response speed is stressed. In a model with starting point variation, the number of categorizations needed will sometimes be greater than and sometimes be less than k_i . If the starting point distribution is unbiased (symmetric), however, k_i can be interpreted as the *average* amount (average across trials) of excess categorizations needed to select perceptual category i .

The relative simplicity of the Poisson random walk led to analytic solutions for its predictions. This is not only advantageous for the technical aspect of fitting the model to data, but also provides a mathematical description of the assumed underlying cognitive processes. For instance, the solution to the problem of trial-by-trial variation in processing rates in combination with the binomial nature of the simple random walk leads to the beta-binomial distribution. This distribution is well known in mathematical statistics as a model for overdispersion (Dean, 1992), that is, a model for data that have a higher variance than could be explained by simpler models with fixed mean-variance relationship (Cox, 1983). For example, the beta-binomial model is appropriate in situations with additional sources of variation, such as trial-to-trial variation in binomial probabilities (Ennis & Bi, 1998).

The beta-binomial distribution describes a general urn model of sampling both with replacement and without replacement (Polyá urn model). Following from the assumptions, the model describes an urn model in which for each ball drawn a ball of the same color is added

⁷ This common assumption may not hold in general. For example, Townsend, Hu, and Evans (1984) report criterion shifts dependent on stimulus confusability in a multidimensional categorization task.

to the urn. In terms of a cognitive processing model, this suggests that a categorization “item x belongs to category i ” increases the probability of collecting evidence for perceptual category i and make the opposite categorization less likely. On the other hand, the longer the process evolves, the lesser is the influence of the initial (beta) distribution parameters—that is, with increasing time the success probability p of the random walk becomes more and more tightly distributed around 0.5. In RT terms, instances of the process that contribute to the higher RT quantiles contain more errors than earlier quantiles. It is this property that lets the model account for errors that are slower than correct responses. Although the effect of trial-by-trial variation on the prediction of slow errors has been known for a long time, the analytic solution allows for an interpretation in terms of a well understood urn model.

Previous studies on detection of briefly presented visual stimuli (e.g., Petersen & Andersen, 2012) have shown that the hazard rates (thus, processing rates) are non-monotonic functions of time. While the exponential function with constant processing rates provides a simple and useful approximation (Bundesen, 1990; Bundesen & Harms, 1999), a detailed analysis revealed that the rates increase after stimulus onset, reach a maximum and decline to a lower steady level thereafter (Petersen & Andersen, 2012). In the Poisson random walk, time-varying processing rates were accounted for by conceiving the input to the random walk as output from a preceding encoding stage that comprises sustained and transient encoding filters. We have demonstrated in Experiment 1 that this stage can be subject to substantial learning, that is, encoding filters can be tuned according to task demands. Time inhomogeneous processing has been used in the diffusion model to account for the deployment of visual attention in spatial cueing tasks (Smith, Ratcliff, & Wolfgang, 2004; Smith & Ratcliff, 2009). If the location of a visual object was correctly cued, the drift rate increased more quickly than if an incorrect location was cued. However, this contrasts with the time inhomogeneous processing assumption in the PRW insofar, as we constrained the processing rates to be proportional functions of time. While this assumption may not hold in

general—as demonstrated by the success of the integrated diffusion model (Smith & Ratcliff, 2009), it may provide the model with sufficient flexibility to predict the leading edge of RT distributions without assuming trial-by-trial variation in non-decision time (see Figure 7). This leads to considerably simpler solutions for obtaining model prediction, so that the full model for two alternatives can be derived analytically.

Another benefit of the simplicity of the Poisson random walk is that it is possible to extend the model to more than two alternatives. This is difficult for random walk models and has only been achieved by using computationally expensive Monte-Carlo simulations (e.g., Leite & Ratcliff, 2010; Usher & McClelland, 2001). One possible extension to any number of perceptual categories is presented and empirically tested in the following section.

A multidimensional random walk model

So far, we have considered the basic model for two perceptual categories (response alternatives). This is probably the most important special case; nevertheless, considering the generality of TVA, it seems desirable to formulate a random walk model for any number of perceptual categories. This generalization can be achieved by formulating the random walk as a competitive network of Poisson counters. This retains all the assumptions of the two-alternative model; the notion that one categorization counts as evidence against all others is now captured by mutual inhibition of the Poisson processes. The only additional assumption to the model with two response categories is a reflecting barrier at zero, so that no counter can ever be inhibited below zero. A schematic representation of the model is given in Figure 10.

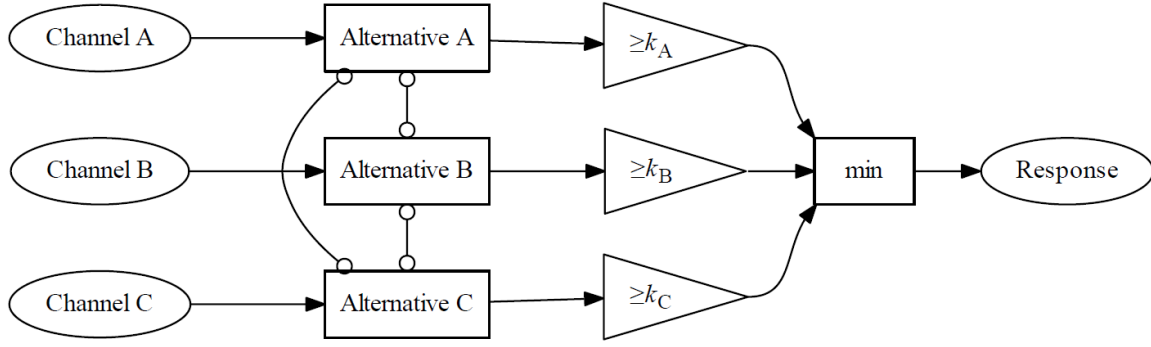


Figure 10. Schematic representation of the Poisson random walk as a competitive accumulator/race model. Arrows represent activation, lines with circles inhibition. The counters receive activation (input) from processing channels and inhibition from all other counters. A response is initiated as soon as the first counter reaches its evidence threshold (k_A , k_B , or k_C).

A response is made as soon as the first counter has obtained a critical amount of evidence. The selection is conceived as a competitive race; if the categorization with most evidence is made again, it progresses towards its criterion, whereas it moves a step towards zero if an alternative categorization is made for x . This reflects a decrease in inhibition, due to which all counters except for the leading one are zero at all times. Only if the process is in the neutral state, that is, all categories have zero evidence, is the network free of inhibition because Poisson counters with zero evidence do not exert inhibition. In this state, whichever tentative categorization for x is made next, the Poisson counter for that category progresses one step towards its evidence threshold. For the sake of clarity in the presentation, we will outline the model for three alternatives, but the generalization readily works for any number of report categories.

Let N_i be defined as the number of steps needed to reach evidence threshold k_i without having reached any $k_i \in \mathbf{k}$ previously. The cumulative distribution function for this event in continuous time t is

$$F_i(t|\mathbf{v}, \mathbf{k}) = \sum_{n=k_i}^{\infty} F(t|n, C) \cdot P(N_i = n|\mathbf{v}, \mathbf{k}) \quad (9)$$

where $\mathbf{k} = (k_1, k_2, k_3)$ are evidence criteria for the three perceptual categories which are processed with rates $\mathbf{v} = (v_1, v_2, v_3)$. The decomposition in Equation (9) is the counterpart to Equation (5) of the two-choice model. It is based on the superposition principle of independent Poisson counters which follows from the memoryless (Markov) property (Townsend & Ashby, 1983, p. 274). Hence, the inter-event time is again exponentially distributed with rate parameter $C = \sum v_i$. Therefore, $F(t|n, C)$ in (9), the waiting time distribution for n exponentially distributed intervals, is an Erlang distribution with shape parameter n and rate parameter C . This part is analogous to the special case of two response categories and easily generalized to three or more alternatives.

The problem is to find an expression for the first-passage probabilities $P(N_i = n|\mathbf{v}, \mathbf{k})$ of the discrete state and discrete step random walk. The discrete nature of the process suggests a formulation in terms of a finite state Markov chain to obtain the first-passage time probabilities. For this approach, one must construct a matrix \mathbf{T} containing the state transition probabilities P_i for each Poisson process. The conditional probability that a tentative categorization for category i (conditioned on the event that a categorization is made at all) is

$$P_i = \frac{v_i}{\sum v_i}. \quad (10)$$

As in the two-choice model, evidence for one category is evidence against all other categories, reflecting inhibition between the counters: At any time, there is maximally one counter active; its activation suppresses activation in all other counters. Without assuming inhibition between the inputs of the counters, this corresponds to a parallel interactive model with inhibition between the counters, but independent inputs (Townsend & Wenger, 2004; Wenger & Townsend, 2006).

All these assumptions are derived from the two-choice model. The only auxiliary assumption for the multi-choice extension of the random walk is that inhibition can never lead to negative activation: any Poisson counter with zero activation maintains zero activation when evidence is sampled in favor of another categorization (Figure 10). It is interesting to note that this assumption can be made for the two-choice model as well, where it has no effect on the model predictions. Although not the only extension possible, we believe that this is the simplest extension of the simple random walk to a multidimensional random walk.

Generalizations based on this auxiliary assumption have been considered for diffusion models as well (Usher & McClelland, 2001; Bogacz 2007; Ditterich, 2010). The simple nature of the discrete state random walk, however, allows for a formulation in terms of a discrete state Markov chain. We refer to Appendix B for details on this approach. In short, we modelled the number of counts of the Poisson counters in the network as a multidimensional process (see Figure 11 for three alternatives).

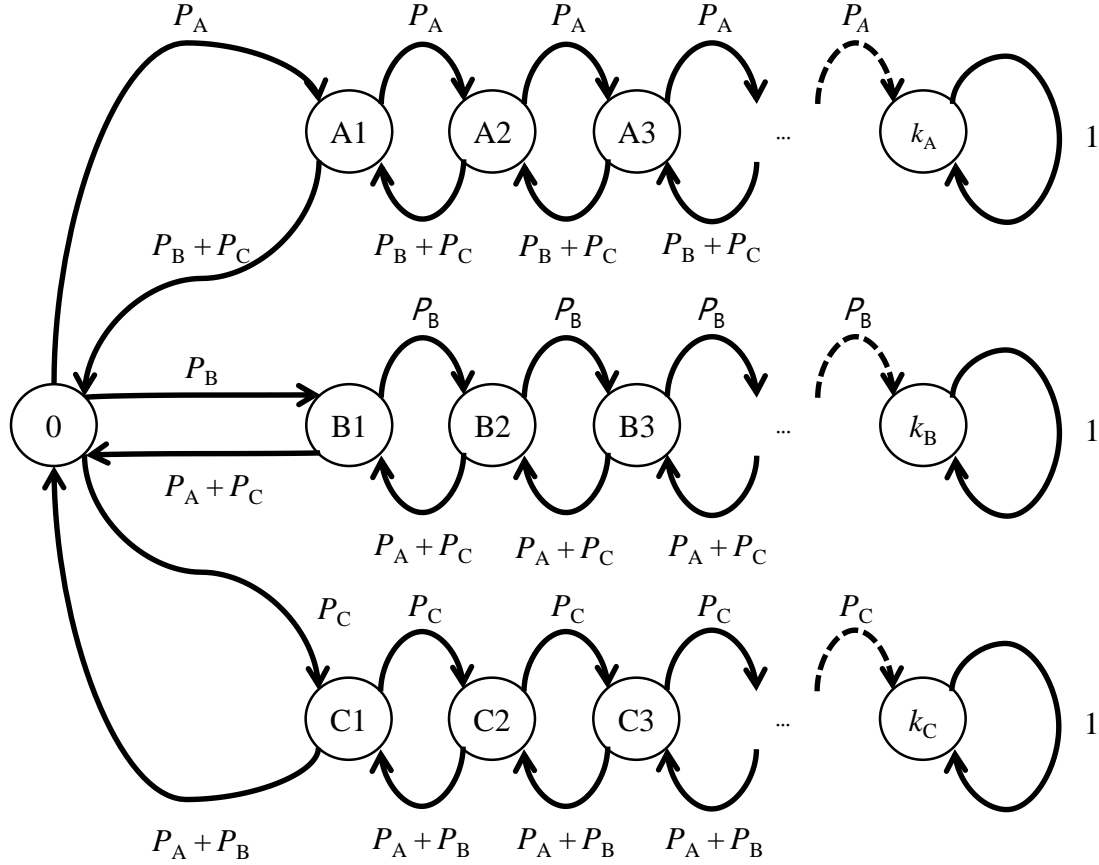


Figure 11. Markov chain of a simple random walk for three response alternatives. A response is initiated as soon as the first counter reaches its threshold (k_A , k_B , or k_C). Transition probabilities P_i are given by Equation (10), note especially that $P_A + P_B + P_C = 1$.

Consider a geometrical representation of a three-alternative process in three-dimensional Euclidian space where evidence for each categorization is represented by the x , y , and z coordinate. Due to competition between counters, the process (Figure 11) can only assume states on the three principal axes because at any time at least two of the coordinates x , y , or z are zero due to inhibition. This is a considerable restriction in states that can possibly be reached compared to, for example, the Poisson counter model in which all counter combinations are permissible (Kyllingsbæk et al., 2012) if not confined by evidence thresholds (Pike, 1973; Townsend & Ashby, 1983). Due to this restriction on permissible states, the multidimensional random walk can be projected from the principal axes onto a line, just as the simple random walk. The only difference is that the state space is non-contiguous if

there are more than two perceptual categories. We have no analytical solution for the first-passage time for such a process, instead, we used standard Markov chain methods to obtain those numerically (Cox & Miller, 1965; Diederich & Busemeyer, 2003). The first passage probabilities $P_i(n|\mathbf{v}, \mathbf{k})$ in (9) were obtained from the state probabilities of states \mathbf{k} after exactly n steps. These need to be related to continuous time by the Erlang distribution (9), just as in the model for two perceptual categories. A version of the model with gamma distributed processing rates (trial-by-trial variation) is also presented in Appendix B.

This generalization of the simple random walk has interesting properties. First, it converges to the solution for two perceptual categories whenever only two categorizations from any number of perceptual categories have non-zero processing rates. Since this may be an empirical question rather than one of a-priori considerations, we consider this an indispensable property. Second, the first-passage time distribution of each perceptual category is approximately a scaled distribution of those of the other categories; the scaling factor is the overall probability of selecting that category. In other words, the conditional first passage time distributions of an unbiased Poisson n -dimensional random walk, conditioned on the event that the respective categorization is selected, are approximately the same for all report categories—at least when confusability is low. This property is well known from unbiased random walk and diffusion models (Laming, 1968; Ratcliff & Smith, 2004; Townsend & Ashby, 1983, p. 305). In the following, we present an empirical test of the Poisson random walk model in a speeded RT experiment with four perceptual categories.

Experiment 2

Methods

We conducted an experiment with a four-alternative speeded categorization task to test the multi-alternative Poisson random walk model. Participants were instructed to categorize Landolt C-rings based on their orientation as fast and as accurately as possible

Participants

Ten new participants (mean age: 23.3 years, range: 19 to 28 years, two males, one left-handed) were tested in Experiment 2. Prior to participation, participants gave their written informed consent. Participants were paid by the hour (DKK 120/approx. 15 US\$). All participants reported normal or corrected to normal visual acuity. The study was approved by the local ethics board (Case no. 2014-01).

Apparatus and stimuli

Target stimuli were dark gray (Weber contrast $C_W = -0.874$) Landolt C-rings with four different orientations. The aperture was either at the top (0), right (90 deg), bottom (180 deg) or left (270 deg). The stimuli subtended 15 deg visual angle and were presented at the center of the visual display. Confusability was manipulated by a spatial jitter of the aperture position: The Landolt rings were randomly rotated by ± 33 deg (easy), ± 39 deg (intermediate), or ± 42 deg (difficult). The stimulus was visible until button press. The apparatus was the same as in Experiment 1 except that we used a four-alternative response box with a circular arrangement of response keys (Cedrus Corporation, San Pedro, CA) to collect the responses. Specifically, there were two response keys on the left and right, and two response keys between them, one closer to the participant (“downwards”) and one farther apart (“upwards”).

Tasks and experimental procedure

Participants were instructed to categorize the Landolt C-rings based on their orientation (i.e., opened rather to the top, left, bottom, or right) and to press a button on the response box accordingly (e.g., pressing the far positioned response key for “upwards” categorizations). Feedback on RT and accuracy was provided after each response. The same inter-trial-interval as in Experiment 1 was used (exponential distribution with mean 1.5 s, plus a fixed interval of 0.5 s). Each participant was tested in 20 blocks; each block comprised 240 trials. We collected 400 responses in each of the $3 \text{ confusability} \times 4 \text{ orientation} = 12$ conditions from each participant—that is, 4,800 trial in total.

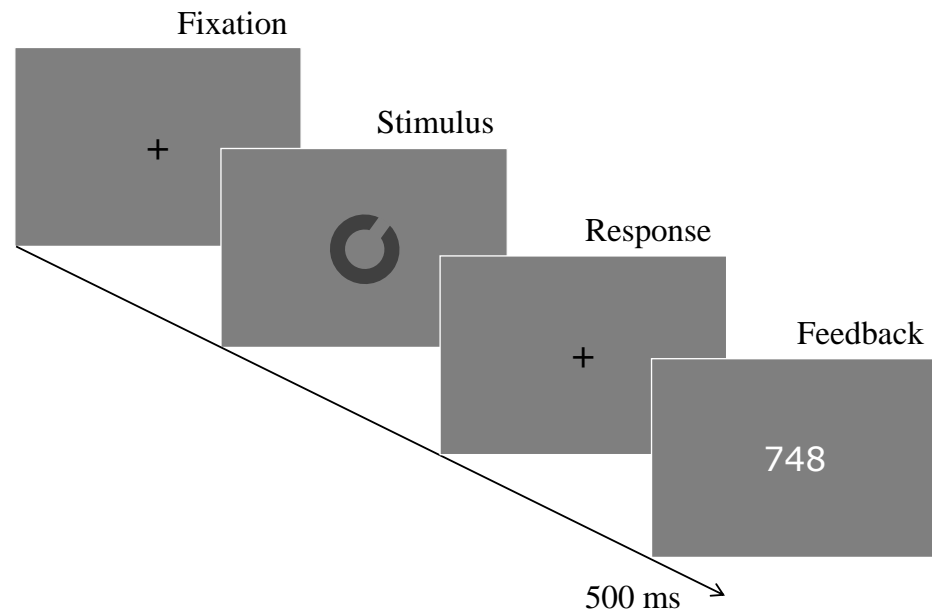


Figure 12. Stimulus protocol in Experiment 2. Feedback about response time (in ms) was given in green (correct responses) or red font (incorrect responses).

Model specification and parameter estimation

We specified a multi-alternative Poisson random walk model with various restrictions on processing rates to explain RT distributions in the 12 conditions. As in the model for two perceptual categories, we assume processing rates $v(x, i)$ of all Poisson processes to be

gamma distributed. Accordingly, processing rates were modelled with a gamma distribution with mean $\bar{v}(x, i) = \eta_{\theta}(x, i) \beta_{\theta}(i) / \xi$ and common scale parameter $1/\xi$. Like in Experiment 1, we modelled trial-by-trial variation on the level of processing rates and assumed that these contained combined variation of perceptual evidence and bias. We restricted perceptual evidence for the correct categorization, $\eta_{\theta}(x, C)$, counterclockwise errors, $\eta_{\theta}(x, E_{-90})$, and clockwise errors, $\eta_{\theta}(x, E_{+90})$, to be the same for all orientations, but varying with confusability (9 parameters). Perceptual evidence for the third possible error $\eta_{\theta}(x, E_{180})$ was set to zero because only very few instances of this type of error occurred (see Results section). Furthermore, we included one bias parameter $\beta_{\theta}(i)$ for each category i ($i = 1, 2, 3, 4$). One of these was a scaling factor; the highest $\beta_{\theta}(i)$ (determined individually) was fixed at unity, resulting in 3 free bias parameters. Additionally, the model included a single evidence threshold (k), the time course of sustained encoding channels (μ_S), and the residual latency parameter T_0 as free parameters. The starting point was fixed to the neutral state (State 0, see Figure 11) for all conditions. Except for perceptual evidence and bias, parameters were the same across all conditions (orientation and confusability). The model had a total of 15 free parameters that were estimated from data (36 distributions represented by a total of 180 quantiles). The same fitting procedure as in Experiment 1 was used—that is, we used quantile maximum likelihood estimation (Heathcote et al., 2002) to fit of the model to the data—that is, the 10%, 30%, 50%, 70% and 90% quantiles of 12 correct RT and 24 error RT distributions obtained in Experiment 2.

Results

Response time and accuracy of report

Response accuracy was about 95% (± 33 deg), 85% (± 39 deg), and 70% (± 42 deg) in the three confusability (difficulty) conditions. Errors occurred almost exclusively as ± 90

deg errors; the number of 180 deg errors was very low (<0.3% for all participants, a total of 39 out of 48,000 responses). Thus, the few observed 180 deg errors were discarded in the model fit by disregarding their contribution to the likelihood and fixing the processing rates for 180 deg errors to zero. The remaining RTs were separated according to orientation and confusability to estimate RT quantiles.

There were significant effects of confusability ($\Delta LL = -1981$, $\chi^2(2) = 3962$, $p < .001$), target orientation ($\Delta LL = -52.78$, $\chi^2(3) = 105.6$, $p < .001$), and their interaction ($\Delta LL = -7.02$, $\chi^2(6) = 14.05$, $p = .029$) on response accuracy. Lower confusability targets (± 33 deg) were detected with significantly higher accuracy than both medium confusability targets (± 39 deg; OR = 3.84, $Z = 32.68$, $p < .001$) and high confusability targets (± 42 deg; OR = 8.32, $Z = 53.74$, $p < .001$). Medium confusability targets were also reported with significantly higher accuracy than highly confusable Landolt rings (OR = 2.17, $Z = 28.44$, $p < .001$). The Landolt rings with aperture oriented to the top were responded to with highest accuracy, significantly higher than rings with rightwards (OR = 1.41, $Z = 9.72$, $p < .001$), leftwards (OR = 1.10, $Z = 2.70$, $p = .007$), or downwards oriented aperture (OR = 1.10, $Z = 2.70$, $p = .007$). Response accuracy was significantly lower in rings with rightwards oriented aperture than in both leftwards (OR = .78, $Z = 7.05$, $p < .001$) and downwards oriented aperture (OR = .78, $Z = 7.05$, $p < .001$). Response accuracy between leftwards and downwards oriented Landolt rings was not significantly different (OR = 1.00, $Z \approx 0$, $p \approx 1$). The interaction between orientation and confusability was such that the effect of direction generally decreased with increasing confusability. For example, the difference in accuracy between rings with aperture oriented to the top vs. right was strongest in the low confusability condition (OR = 1.86) and smaller in medium confusability (OR = 1.41) as well as high confusability (OR = 1.31). The other differences show a similar pattern.

In terms of mean RT, the effect of confusability was significant ($F_{2,99} = 46.52$, $p < .001$), whereas neither the effect of target orientation ($F_{3,99} = 0.43$, $p = .735$) nor their

interaction was significant ($F_{6,99} = 0.12, p = .994$). Mean RT was significantly lower in the lowest confusability condition than in the medium ($t_{108} = -4.99, p < .001$) and the high confusability condition ($t_{108} = -9.08, p < .001$). These differences were 80.9 ms (95%-CI: 49.1 to 112.6 ms) and 161.7 ms (129.9 to 193.5 ms), respectively. Finally, responses to medium confusability targets were significantly faster (80.8 ms, 49.1 to 112.6 ms) than were those to high confusability targets ($t_{108} = -4.99, p < .001$). In sum, we had quite a regular pattern of slowdown from low to high confusability with responses to targets of intermediate confusability lying quite halfway in between. No effects of stimulus orientation on response speed were observed. Concerning accuracy, however, the pattern was more complicated. There seems to have been a bias towards rings with aperture oriented upwards and against rings oriented rightward. Otherwise, accuracy deteriorated from low to intermediate confusability but even more so for highly confusable targets. These quite different patterns (RT, accuracy) need to be accounted for by the Poisson random walk model.

Fit of the n-choice Poisson random walk model

The Poisson random walk model could explain the distributions of correct responses and all the different types of erroneous reports (Figure 13). The overall goodness of fit was excellent, the model accounted for 93.0 % (89.7 to 98.5%) of variation. There is a close agreement between model and data for correct responses and the distributions of incorrect responses are also well predicted overall; notable differences between model and data were only observed in some error RT distributions in the lowest confusability condition. However, these distributions contained by far the least amount of responses (less than 5% of responses, or about 24 RTs per participant) so that few contaminants can already distort the density estimates of those distributions. The model fit is less affected by these contaminants since the model fit was based on robust estimators (distribution quantiles, see Methods section above).

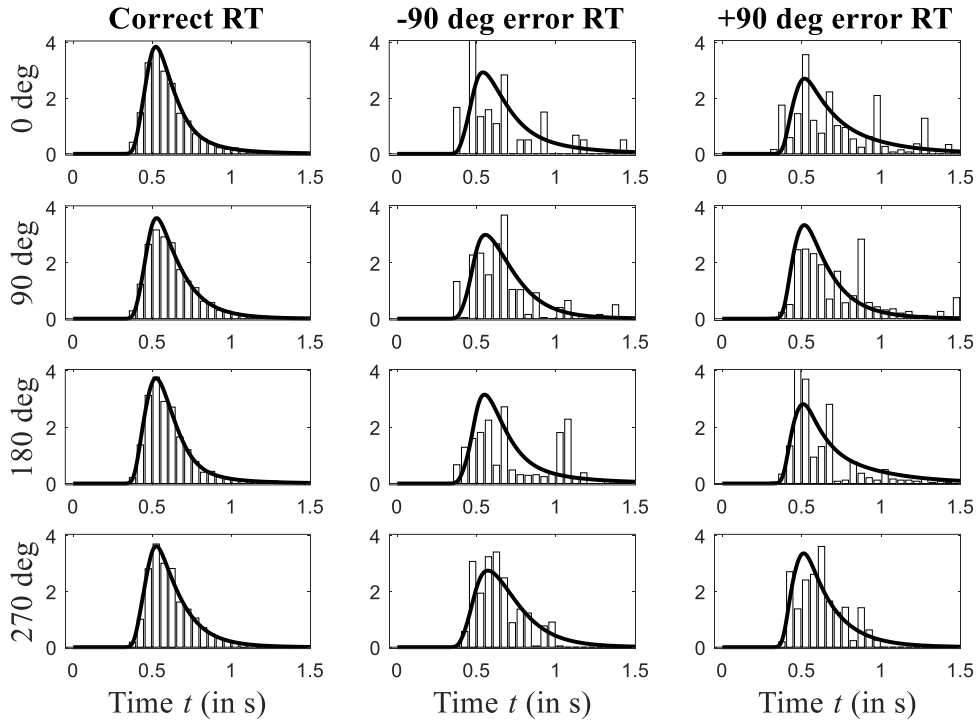
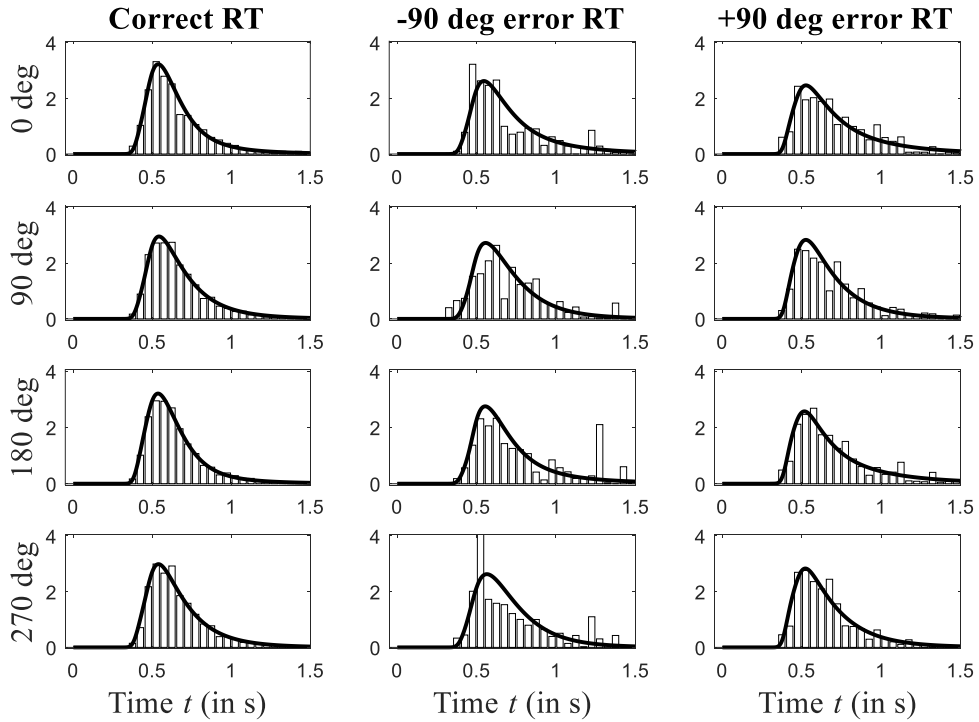
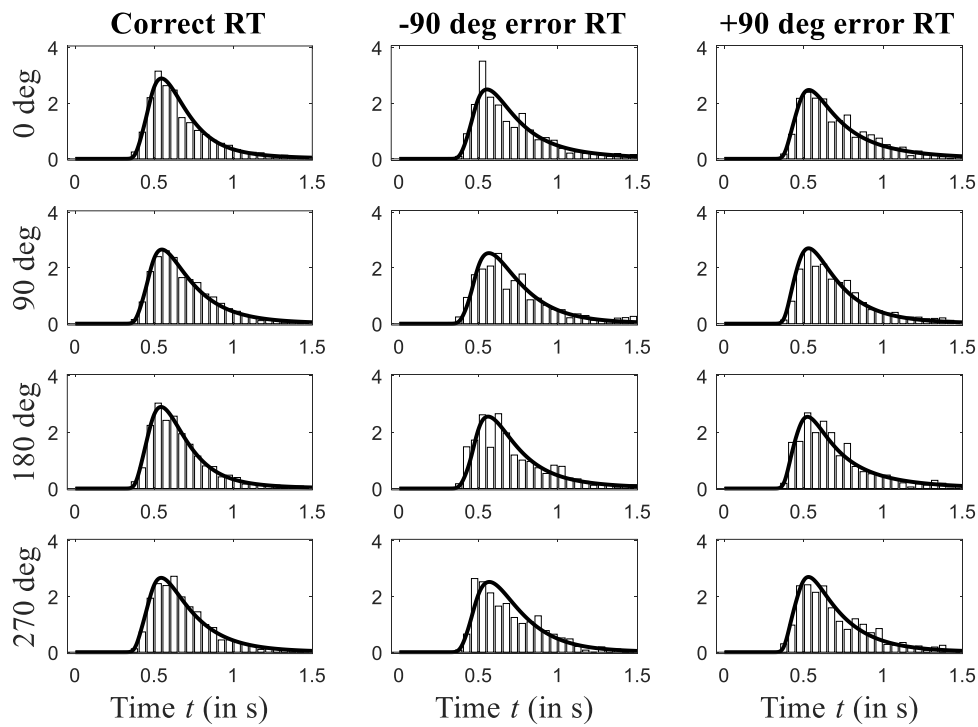
Confusability: ± 33 deg**Confusability: ± 39 deg**

Figure 13. Empirical RT density estimate (histograms) and model prediction (unmarked points connected with lines) for group averaged data ($N = 10$) of Experiment 2. The panels in the first column are data and predictions for correct RT; the two remaining columns are -90 (i.e., counterclockwise) and $+90$ deg (i.e., clockwise) confusions. Panels in each row show responses to differently oriented stimuli (0, 90 deg, 180 deg, 270 deg, respectively).

Confusability: ± 42 deg*Figure 13—continued*

The confusion matrix (Table 5) containing the overall probabilities of reporting a perceptual category given stimulus x was also closely matched by the model. The minor differences between data and predictions are mainly caused by the restriction of perceptual evidence for correct and incorrect categorizations being identical for all stimulus orientations. Stated differently, this specification did not allow for an interaction of confusability and orientation in the processing rates.

As in the model with two report categories, we also fitted a “full” model in which all perceptual evidence parameters were free to vary. Due to the addition of no less than 24 parameters (33 perceptual evidence parameters, instead of 9), the model fit improved further ($R^2 = 0.945$, range 0.921 to 0.995), that is, an average of 1.5% (0.8 to 2.7%) of additional variation was explained by the full model. Since this model can account for an interaction between orientation and confusability, the difference in goodness-of-fit was significant in all

participants (all $\chi^2 > 40.3$, $df = 24$, $p < .02$). Despite this further improvement in model fit, we present the restricted model without those additional parameters.

Table 5.

Observed (predicted) group average confusion matrix in Experiment 2.

Stimulus (x)		Probability of reporting category			
		up	down	left	right
± 33 deg	up	.96 (.95)	<.001 (0)	.01 (.02)	.03 (.03)
	down	<.001 (0)	.94 (.95)	.03 (.04)	.02 (.02)
	left	.02 (.03)	.02 (.03)	.95 (.94)	<.001 (0)
	right	.02 (.02)	.05 (.04)	<.001 (0)	.93 (.94)
± 39 deg	up	.84 (.84)	<.001 (0)	.05 (.06)	.10 (.10)
	down	.002 (0)	.83 (.83)	.10 (.11)	.07 (.06)
	left	.09 (.10)	.07 (.08)	.83 (.82)	<.001 (0)
	right	.07 (.07)	.13 (.12)	<.001 (0)	.79 (.81)
± 42 deg	up	.71 (.72)	.001 (0)	.11 (.11)	.18 (.17)
	down	.002 (0)	.70 (.71)	.18 (.18)	.12 (.11)
	left	.16 (.18)	.15 (.15)	.69 (.68)	<.001 (0)
	right	.13 (.14)	.22 (.21)	.001 (0)	.65 (.66)

Note—Values in parentheses are model predictions. Proportions (probabilities) printed in bold are for correct reports.

As previously described, we specified the processing rates similar to common applications of TVA. According to TVA, the processing rates in single stimulus recognition are composed of perceptual evidence and bias. The processing rates were obtained according to Equation (7) with perceptual evidence parameter $\eta_\theta(x, i)$ for the categorization “item x belongs to category i ” and bias parameter $\beta_\theta(i)$ (here: $i = 1:4$, for 0, 90, 180, and 270 deg). The bias parameters were scaled to lie between zero and unity so for each participant only three $\beta_\theta(i)$ were free parameters. Perceptual evidence parameters were constrained so that all correct categorizations (C), all -90 deg errors (E_{-90} , i.e., counterclockwise erroneous categorizations) and all $+90$ deg errors (E_{+90} , i.e., clockwise erroneous categorizations) were constant across the four directions but different across stimulus confusability conditions.

The processing rates for correct categorizations decrease with increasing confusability while those for errors increase. The increase in processing rate for correct categorizations is

significant ($F_{2,18} = 9.143, p = .002$). The processing rate for correct categorizations of low confusability targets is significantly higher than that of highly confusable targets ($t_{18} = 4.157, p < .001$) and significantly higher than that of targets of intermediate confusability ($t_{18} = 2.946, p = .009$). Processing rates for intermediate confusability and high confusability did not differ significantly ($t_{18} = 1.212, p = .241$). The difference in processing rates for correct categorizations between low and medium confusability was 6.3 Hz (95%-CI: 2.1 to 10.4 Hz), between low and high confusability 8.8 Hz (4.7 to 13.0 Hz), and between medium and high confusability 2.6 Hz (−1.6 to 6.7 Hz). In the most difficult condition, the correct categorization for x has a processing rate of less than one half of the overall processing capacity (cf. Table 6); nonetheless, it is higher than all other categorizations and is therefore selected most often (cf. first row in Figure 13).

The processing rates of 90 deg counterclockwise errors also significantly differed across confusability conditions ($F_{2,18} = 92.07, p < .001$): Low confusability targets had significantly lower processing rates for 90 deg counterclockwise errors than targets of intermediate confusability ($\Delta v = -7.8$ Hz, $t_{18} = -6.88, p < .001$) and highly confusable targets ($\Delta v = -15.4$ Hz, $t_{18} = -13.57, p < .001$). Intermediate confusability targets had a significantly lower processing rate for those errors than targets of high confusability ($\Delta v = -7.6$ Hz, $t_{18} = -6.69, p < .001$). The identical pattern was obtained for the processing rates of 90 deg clockwise erroneous categorizations ($F_{2,18} = 54.79, p < .001$). Processing rates for those errors were significantly lower for low confusability targets than targets of intermediate ($\Delta v = -9.5$ Hz, $t_{18} = -5.69, p < .001$) and high confusability targets ($\Delta v = -17.4$ Hz, $t_{18} = -10.45, p < .001$); intermediate confusability targets had significantly lower processing rates for 90 deg clockwise errors than high confusability targets ($\Delta v = -7.9$ Hz, $t_{18} = -4.76, p < .001$). The processing rates of stimuli of different orientation were not significantly different ($F_{3,27} = 0.62, p = .608$). Thus, in line with the RT results but contrary to accuracy, we did not obtain evidence for perceptual bias in the estimated processing rates of the model. We did not

test for any interactions between confusability and orientation because we did not specify such an interaction in the processing rates of the model. The mean processing rates $\bar{v}(x, i)$, expanded for all stimuli x and response categories i , are summarized in Table 6 and the model parameters in Table 7.

Table 6.

Group average of estimated mean processing rates $\bar{v}(x, i)$ in Experiment 2.

Stimulus (x)		Processing rates for categorizations (i)			
		up	down	left	right
± 33 deg	up	70.4	(0)	17.0	20.3
	down	(0)	71.0	21.1	16.6
	left	21.2	17.9	67.6	(0)
	right	17.5	21.6	(0)	66.7
± 39 deg	up	63.8	(0)	24.8	29.3
	down	(0)	64.4	30.5	24.1
	left	30.8	25.9	61.6	(0)
	right	25.5	31.2	(0)	59.8
± 42 deg	up	61.0	(0)	32.3	36.9
	down	(0)	61.8	38.4	31.4
	left	38.9	33.9	59.2	(0)
	right	33.2	39.5	(0)	57.3

Note—Four report categories ($i = 1, 2, 3, 4$) and 12 different stimuli (x) were used in Experiment 2. Thus, the model contains 48 processing rates as parameters (not free parameters). Values printed in bold are processing rates of correct categorizations, values in parentheses are fixed. With various restrictions, all processing rates were obtained from 12 free parameters (see Table 7 and text for details).

The average response threshold was $k = 3.4$ for all perceptual categories. This estimate is slightly lower than in Experiment 1, however, in contrast to Experiment 1, we did not allow for starting point variation. The estimate of $1/\xi = 2.1$ suggests that the variance of processing rates was about twice their expected value. The sustained encoding channels (μ_S) responded slightly slower than in Experiment 1 while the estimated residual latencies (Experiment 2: 237 ms, Experiment 1: 239 ms) are almost identical. We also fitted a time homogeneous Poisson model with trial-by-trial variation in non-decision time. As already observed in

Experiment 1, this model fitted approximately as well as the time inhomogeneous model with constant non-decision time ($R^2 = 0.930$, range 0.897 to 0.985).

Table 7.

Parameter estimates of the 4-report category data of Experiment 2.

Parameter		Group average ($\pm SD$)
$\beta_\theta(i)$	up	0.93 (± 0.07)
	down	0.94 (± 0.10)
	left	0.90 (± 0.10)
	right	0.88 (± 0.07)
$\eta_\theta(x, C)$	33 deg	75.4 (± 23.0)
	39 deg	68.4 (± 19.3)
	42 deg	65.5 (± 18.5)
$\eta_\theta(x, E_{-90})$	33 deg	18.9 (± 4.6)
	39 deg	27.5 (± 5.4)
	42 deg	35.8 (± 8.4)
$\eta_\theta(x, E_{+90})$	33 deg	22.9 (± 10.2)
	39 deg	33.2 (± 13.6)
	42 deg	41.9 (± 16.4)
$1/\xi$		2.09 (± 1.89)
k		3.40 (± 0.52)
μ_S		18.0 (± 6.6)
T_0 (in s)		0.237 (± 0.057)

Note—Bias parameter $\beta_\theta(i) \in [0, 1]$. Perceptual evidence for correct categorization ($\eta_\theta(x, C)$) and errors ($\eta_\theta(x, E_{-90})$, $\eta_\theta(x, E_{+90})$) are the same across orientations. Processing rate is gamma distributed with scale $1/\xi$. The four evidence criteria k_i are placed at k , sustained encoding channels are described by μ_S . Parameter T_0 is the residual latency. The standard deviation of parameters (in parentheses) was obtained across participants.

For the sake of comparability, we also report quantile probability plots as in Experiment 1 with two perceptual categories. For these plots, we pooled data and model predictions of all errors and calculated the predicted conditional distributions which were then averaged across orientation conditions. This was done to avoid cluttering the quantile probability plot with too many distributions that are located at more or less the same position. The quantile probability plot shows the similarity with two-choice random walk models of response times, for example the diffusion model (Ratcliff, 1978). When confusability is low,

and the correct categorization is easily found, both errors and correct responses are faster than responses to highly confusable stimuli where the correct categorization is difficult to find.

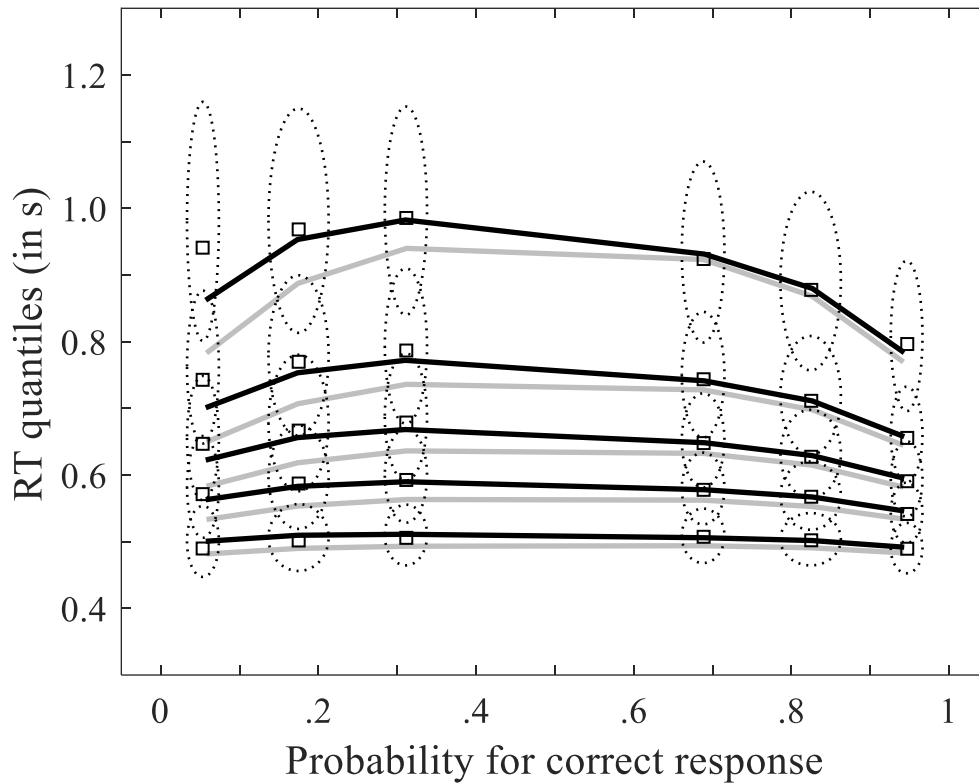


Figure 14. Data and model predictions of Experiment 2. The three rightmost distributions (10%, 30%, 50%, 70% and 90% quantiles, from bottom to top) are distributions of correct responses, the three leftmost distributions are combined ± 90 deg errors at each level of confusability (± 33 deg; ± 39 deg; ± 42 deg orientation jitter). The unmarked points connected with lines are model predictions of the model with trial-to-trial variability (black) and without this assumption (gray); squares are data. The dotted ellipses represent 95% confidence intervals for response probabilities and distribution quantiles (both determined by bootstrap).

Within each difficulty condition, errors were relatively slow compared to correct responses (Figure 14). As expected, the model could capture this aspect of the data by the trial-by-trial variation in processing rates. In contrast to the two-alternative categorization data (Experiment 1), errors faster than correct responses were rare, and we therefore fitted the model with a constant starting point ($z = 0$). As a result, the predictions for the 10% distribution quantiles almost constant across conditions (Figure 14). The excellent fit to the

correct RT distributions is apparent; model predictions for the incorrect RT distributions also well match the observed data. The model fit to the incorrect RT distributions of intermediate and difficult stimuli was excellent, only in the easiest condition there is a slight mismatch between model and data in the highest quantiles (70% and 90% distribution quantiles). Clearly, estimates of these quantiles varied most because only few errors (less than 5%) occurred in this condition.

Finally, we compared the multi-alternative Poisson random walk with and without the assumption of trial-by-trial variability in processing rates and found that without this variation the goodness-of-fit deteriorated, especially in the incorrect RT distributions (Figure 14). This restriction eliminates a single parameter (ξ) and the accompanying lack-of-fit relatively was low in terms of explained variation ($R^2 = 0.925$, range 0.885 to 0.984), that is, an average of 0.5% less explained variation. However, some predictions of the restricted model lie outside the 95% confidence intervals around the data points, so we discarded this model. Thus, trial-by-trial variation of processing rates is important to account for shapes of correct and error RT distributions in categorization with multiple report categories, as it is with two report categories. Without this source of variability, the predicted condition distributions are restricted to approximately equal distributions of errors and correct RT. Due to the assumption of mutual inhibition and non-negative accumulators, the model predicts slightly slower errors in high confusability conditions, whereas in low confusability conditions predictions for error and correct RTs are practically the same (Figure 14).

Discussion

With a single and simple additional assumption, it is possible to generalize the Poisson random walk model to any number of perceptual categories. Despite its simplicity, the extension turned out to account well for data obtained in an experiment with multiple report categories. We consider it a random walk model in the sense that each categorization “item x

belongs to category i ” is simultaneously evidence against all other categories j . The multi-alternative extension is related to the leaky competitive accumulator model (LCA, Usher & McClelland, 2001). The LCA is a model of leaky competitive Gaussian accumulators (Ornstein-Uhlenbeck processes) with mutual inhibition and decay, but in contrast to the LCA model we do not consider leakage at this point. In addition, the inhibition weights in the PRW are fixed to unity—that is, a tentative categorization for one category counts as unit evidence against all other categories.

Consequently, the interpretation of the model as a cognitive model for categorization of confusable stimuli is largely the same as in the model for two perceptual categories. Only the interpretation of relative evidence thresholds changes somewhat in the model for n perceptual categories: In the Poisson random walk for two perceptual categories, an evidence threshold of k (and $z = 0$) means that either of the two categories needs exactly k tentative categorizations more than the other. In the multivariate random walk, this changes to the requirement that category i needs *at most* k_i categorizations more than the category with second most evidence. In situations with highly confusable stimuli—that is, several categorizations for x have a similar processing rate, it will often be the case that there are several categorizations with a similar amount of evidence. Then, mutual inhibition will be significant, and, in many instances, the process will frequently return to the neural state. This is exactly the scenario where full inhibition, that is, inhibition that leads to negative activation (corresponding to the notion of the evidence threshold that category i needs exactly k_i more tentative categorizations than *all* others) will lead to a dead network, that is, one where inhibition overweighs activation and the process has a non-zero probability of never reaching any evidence criterion.

In the Poisson random walk, as in diffusion-based models (Ditterich, 2006; Usher & McClelland, 2001), the problem of a dead network is solved by the additional assumption that inhibition may never lead to negative activation. With this assumption, the eventually selected

category i in a highly contested race as described above will often need less than k_i tentative categorizations more than the categorization with second most evidence. This ensures that the process is well founded and that it will reach one evidence threshold eventually; however, the interpretation of evidence thresholds becomes less straightforward compared to the special case of two report categories. We do not consider this to be a critical weakness, though, as an interpretation of evidence thresholds in any case is meaningful only with a constant starting point or in a comparison of different conditions all with the same starting point distribution (e.g., speed-accuracy trade-off, see Appendix D).

As a rather extreme example, consider a situation where a stimulus is to be categorized as belonging to either one of three categories ($k = k_1 = k_2 = k_3$). Suppose that the first ten tentative categorizations are five each for the first two categories and that they occur in a way that neither evidence criterion is reached (e.g., when they are strictly alternating). Then, these ten tentative categorizations cancel each other out completely, so that the process is in the neural state after these ten tentative categorizations have been made. The third alternative then only needs k tentative categorizations to be selected. If this happens, the eventually selected categorization (Alternative 3) may have even *less* evidence than any of the other perceptual categories (i.e., if $k < 5$). This is because the individual counts are not represented but lost due to inhibition. As stated above, this is a rather extreme example, mostly relevant in situations of high confusability conditions. The inhibition assumption is supported at least by the present data, as suggested by the overall fit of the model and its fit to the highest confusability condition, which is particularly well predicted. It remains to be seen whether this assumption is still plausible if the input changes over time (Ratcliff, 1980), for example, in a redundant signals task in which the redundant targets are presented with onset asynchrony (Blurton, Greenlee, Gondan, 2014; Miller, 1986; Schwarz, 1989).

In general, when stimulus confusability is high so that evidence is frequently obtained for different categories, the model works similar to a runs model (Audley, 1960; Audley &

Pike, 1965; Ratcliff, 1988) according to which selection is based on evidence accumulated successively. In a runs model, perceptual category A , say, may be chosen as soon as k tentative categorizations in favor of A are obtained. It is due to this selection rule that the predicted distribution quantiles for incorrect responses are slightly higher than are those of correct responses (Figure 14). The runs model is related to the counter model (Audley & Pike, 1965) which generally predicts errors to be slower than correct responses (Ratcliff & Smith, 2004; Townsend & Ashby, 1983).

Although we consider the presented extension from binary to multiple choice a well-motivated and rather simple extension, there are, clearly, many other possibilities to achieve this goal. One of these was offered by Zhang and Bogacz (2010); it is also based on Poisson counters and the assumption of mutual inhibition between these counters. This model is conceived as a model for neural spike trains and is more general because it assumes populations of Poisson counters (instead of one per alternative). Although the model is restricted by the assumption that all counters for incorrect choices have the same processing rate (making it unsuitable for RT modelling unless errors are pooled), simulations are required to obtain predictions of that model (Zhang & Bogacz, 2010). Other possibilities to extend the Poisson random walk from 2 to n alternatives include different weights of inhibition (with a lower inhibition the reflecting barrier may be avoided), or the assumption that evidence in favor of each alternative form a simplex with the total evidence always sum to a constant. Note that the extensions above are different to those developed for the diffusion model to account for continuous report categories (Kvam, 2019a, 2019b; Ratcliff, 2018; Smith, 2016) in that they describe evidence accumulation for n discrete dimensions. A general framework for higher dimensional models was laid out by Kvam (2019a), and the hypersphere diffusion model recently developed by Smith and Corbett (2019) is an instance of a higher dimension model for continuous report categories.

One of the challenges with models such as the Poisson random walk for n perceptual categories is to find reasonable constraints on the processing rates $v(x, i, t)$ (Smith & Van Zandt, 2000). A model without any constraints on these will have too many parameters and be hard to falsify (but see Ratcliff, 2002). On the other hand, RT models are restricted mostly by the requirement to simultaneously explain distributions of correct responses *and* errors. Thus, the additional information obtained from differentiating between different kinds of errors constrains the model since their all distributions need to be accounted for simultaneously. Then, fitting an n -choice model can be more restrictive than the two-choice model, depending on the restrictions placed on the processing rates. In the applications presented here, the two-choice model (Experiment 1) and the 4-choice model (Experiment 2) had a similar number of parameters (13 and 15, respectively) and a comparable goodness-of-fit across participants (Experiment 1: 97.6%, Experiment 2: 93.0% explained variation). But the two-choice model explained 12 RT distributions, whereas the 4-choice model explained no less than 36 distributions. In general, data from experiments with multiple report categories can potentially restrict RT models more than data with only two report categories because the number of distributions to be explained grows faster than the number of additional parameters. To illustrate this point, we present an application to speeded RT data obtained in an experiment on absolute identification with ten perceptual categories (Lacouture & Marley, 2004) in Appendix E.

General discussion

We presented and tested a Poisson random walk model of speeded categorization of confusable stimuli. The proposed model is based on the theory of visual attention (Bundesen, 1990) and thereby theoretically founded in similarity-choice theory (Luce, 1963; Shepard, 1957). It can be conceived as a sequential sampling extension of the similarity-choice rule in which evidence is sampled until a relative response criterion is reached. These assumptions

lead to a random walk: At each step of the random walk, a tentative categorization “item x belongs to category i ” is made with a probability derived from the Shepard-Luce choice rule. The extension to a random walk allows for response selection under time pressure in presence of decision noise. Decision noise arises with stimulus confusability—that is, when a stimulus activates several perceptual categories. Assuming exponentially distributed intervals between any two tentative categorizations, the number of categorizations after time t follows a Poisson distribution, hence the name. In this regard the Poisson random walk model is similar to the Poisson counter model described by Pike (1973, see also Townsend & Ashby, 1983; Van Zandt, Colonius & Proctor, 2000) with the sole difference that the selection of a category is based on a relative evidence threshold (Smith & Ratcliff, 2004), rather than on absolute thresholds.

Our choice of relative evidence thresholds was motivated on the one hand by the previous success of models with relative evidence thresholds (Ratcliff & Smith, 2004; Ratcliff, 1981; Ratcliff, 1988; Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008; but see Van Zandt, Colonius & Proctor, 2000). On the other hand, it was motivated by the fact that models implementing the sequential probability ratio test provide an account for optimal performance (i.e., maximizing accuracy) under time constraints (Bogacz, Moehlis, Holmes, & Cohen, 2006; Bogacz, 2007; Zhang & Bogacz, 2010). It has become clear, however, that an empirical distinction between those two model classes can be difficult (Khodadadi & Townsend, 2015) or impossible for some models (Jones & Dzhafarov, 2014). Under restrictive assumptions, one of those being constant response thresholds, one could compare predictions from the Poisson random walk and the Poisson counter model to directly address the question whether data is better described by absolute or relative evidence because the two models only differ in this single, though fundamental assumption. Clearly, the results of such a comparison depend on the validity of the assumptions; for example, if response thresholds vary over time, the Poisson counter model can make predictions identical to those of a

random walk model (Khodadadi & Townsend, 2015). Nevertheless, additional dependent measures, such as confidence ratings (Pleskac & Busemeyer, 2010), could help to distinguish between the two model classes.

The close relationship between the Poisson random walk and similarity-choice models based on the Shepard-Luce choice rule leads to close connections with other models based on similarity-choice, such as the generalized context model (Nosofsky, 1986), the CODE theory of visual attention (Logan, 1996), the executive control theory of visual attention (Logan & Gordon, 2001), the instance theory of attention and memory (Logan, 2002), and, of course, the theory of visual attention (Bundesen, 1990). In this view, the Poisson random walk serves as a response selection mechanism while the respective models provide input to the random walk in form of its processing rates. This approach has been suggested by Logan (1996) who used the Poisson counter model for response selection.

Like other RT models based on diffusion processes (Diederich, 1997; Ratcliff, 1978, 1981; Ratcliff & Rouder, 1998; Smith & Ratcliff, 2009; Usher & McClelland, 2001) or accumulators (Brown & Heathcote, 2005, 2008), the Poisson random walk model does not only predict mean RT, but whole distributions of correct and incorrect responses. The flexibility required to fit RT distributions is achieved by common trial-by-trial variation assumptions (Brown & Heathcote, 2005, 2008; Ratcliff, 1978; Ratcliff & Rouder, 1998). Although response times have been used to estimate TVA parameters (Bundesen, 1990), they have received much less attention than accuracy of report. With this paper, we want to change this state of affairs by proposing a TVA-based response time model that so far has proven to be an excellent account of RT distributions. The model is simple and its parameters are tightly connected to cognitive processes. Its simplicity allows for an analytic solution in the special case of binary perceptual decisions and efficient numerical evaluation in the general case of three or more report categories. This generalization is particularly intriguing and matches the generality of TVA, but for response times. In the future, many more applications to classic

attention-based RT tasks are conceivable, such as spatial cueing (e.g., Posner et al., 1980), selective attention (Eriksen & Eriksen, 1974) or visual search (e.g., Treisman & Gelade, 1980).

Discrete state model

The simplicity of the model is largely due to its discrete state space—that is, the number of excess tentative categorizations in favor of one perceptual category. This seems at odds with other models in the field that are based on a continuous state space (Link & Heath, 1975; Ratcliff, 1978; but see Nosofsky & Palmeri, 1997a for another discrete state model). This distinction, however, is more a technical than a theoretical one. It has long been known that most discrete and continuous models are formally related (*Grice-representability*, Dzhamfarov, 1993). As a consequence, none of these models is uniformly superior, so the question is rather one of mathematical tractability and personal preference.

This has also been pointed out in a comparison of the diffusion model with the Poisson counter model (Van Zandt, Colonius, & Proctor, 2000). The *technical* issue with fitting a model with discrete state space is addressed in that study as well. Van Zandt et al. (2000) used a different approach than we did in this study: they performed a grid search on plausible threshold values first and then used the best result in a subsequent estimation step for the remaining parameters. Here, we used a probability mixing approach to allow the thresholds to take real values and perform parameter estimation for all parameters simultaneously (Appendix C). This approach may not work optimally with multiple integer-valued parameters, which is why care was taken to avoid more discrete valued parameters—for example, the starting point was either fixed or described by a starting point distribution with only real parameters (i.e., the beta-binomial distribution).

Relation to other models

The closest relative of the Poisson random walk model is Nosofsky and Palmeri's (1997a) exemplar-based random walk model (EBRW) of speeded classification. The EBRW is a combination of Nosofsky's (1986) Generalized Context Model (GCM) and Logan's (1988) Instance Model (Palmeri, 1997). As in the Poisson random walk model, it is assumed that several classifications are made successively to optimize performance under time pressure. Instead of categorizations of the type "item x belongs to category i " (as in TVA), the GCM models classifications of the type "object x is a member of class i ". The similarity is apparent, but the processing rates are obtained in an entirely different manner. The GCM regards class i to be represented by exemplars located in a psychological space; the closer two instances are located to each other, the more similar they are (i.e., a metric). The assumptions of the EBRW for classification can be obtained by inserting the processing rates obtained from the Generalized Context Model (GCM, Nosofsky, 1986) into the Poisson random walk:

$$v_{ij} = M_j \exp(-cd_{ij})$$

where d_{ij} is the distance between exemplar i and j in psychological space, c is a sensitivity (or scaling parameter), and M_j is the memory strength for exemplar j . Without further adjustments these rates can be used as processing rates in the Poisson random walk. For example, the probability to classify item i into Category J is

$$P(J|i) = \frac{\sum_{j \in J} v_{ij}}{\sum_K \sum_{k \in K} v_{ik}}$$

that is, the sum of all processing rates v_{ij} for Category J divided by the sum of processing rates of all categories—the processing capacity C . In the random walk, this is the probability of taking a step towards the evidence threshold for Category J .

There are only minor differences between the EBRW and the Poisson random walk. One of these is that the time to retrieve an exemplar and increase the counters is explicitly included in the EBRW. This is achieved by modelling the interstep time in the process not as an exponential, but as a shifted exponential distribution. In other words, the non-decision time (T_0) is made dependent on the number of steps. This assumption can easily be included in the Poisson random walk as well. In fact, we added this parameter but found that the model fit increased only negligibly and that the parameter was close to zero. In the model fits presented, the time to increase counters was therefore fixed at zero. An obvious difference between the EBRW and the Poisson random walk is that the latter predicts RT distributions, not only for two-choice data but also in n -choice tasks. Both the EBRW (Nosofsky & Palmeri, 1997a) and the Poisson random walk presented in this paper contain their predecessors as a special case: When the evidence thresholds are set to unity, the first classification made determines the response and the EBRW becomes the general context model (Nosofsky, 1986). The same holds for the Poisson random walk that becomes TVA, in which the first categorization that finishes processing is the one that is reported (Bundesen, 1990). Essentially, both the EBRW and the Poisson random walk are extensions in which repeated classifications (GCM) or categorizations (TVA) are made until a pre-set evidence threshold—possibly greater than unity—is reached. Seen this way, the EBRW and Poisson random walk can be conceived as response selection mechanisms for the underlying cognitive models.

Another model that can easily be combined with the Poisson random walk is the CODE Theory of Visual Attention (CTVA, Logan, 1996). The CTVA model is in turn a combination of TVA and the contour detector theory (CODE, van Oeffelen & Vos, 1982). It extends TVA from a purely object-based attention theory to a spatial attention theory. According to CTVA, object features can be represented by distributions of stimulus evidence across space. Items are grouped by application of a threshold on that distribution. Depending on that threshold, proximately located objects or more distantly located objects may be

grouped together. Through this grouping, confusability may arise because the features of all grouped items are processed together (“feature catch”). The feature catch can be easily integrated into TVA (Logan, 1996) and hence into the Poisson random walk. To this end, the processing rates are conceived as a combination of perceptual evidence strength, perceptual bias and the feature catch parameter c :

$$v(x, i) = c_x \eta(x, i) \beta_i \frac{w_x}{\sum_{z \in S} w_z}$$

with attentional weights

$$w_x = \sum_{j \in R} c_x \eta(x, j) \pi_j.$$

The CODE theory of visual attention was initially proposed as a counter model; thus, it is straightforward to combine it with the Poisson random walk model, since only the evidence thresholds are changed (relative vs. absolute thresholds). The combination of a spatial attention theory with the RT theory could be useful for explaining distance and grouping effects in experiments on the flanker task (Eriksen & Eriksen, 1974) or in visual search (e.g., Banks & Prinzmetal, 1976).

In general, the Poisson random walk model is well compatible with theories based on the similarity-choice rule (Luce, 1963; Shepard, 1957; see Logan, 2004, for a review of potential candidates). These include the executive control theory of visual attention (ECTVA, Logan & Gordon, 2001) and the instance memory of attention and memory (ITAM, Logan, 2002). Covering attention and memory processes, the scope of ITAM itself is already impressive; the combination with the Poisson random walk to explain RT distributions in tasks with multiple report categories is certainly intriguing. Yet, models that are not based on similarity choice could also be combined with the Poisson random walk. The discrete nature of the random walk links elegantly to theories on visual search, which usually assume some

serial processing of elements in the display if the search is inefficient because the display must be scrutinized for the presence of a target. If one assumes that the serial search is performed on single items (Wolfe, 2007), this would result in a discrete step visual search (Moran, Zehetleitner, Müller, Usher, 2013). This notion is not restricted to a single item being processed at a time. One theory that conceives search in this way is Guided Search 4.0 (GS4, Wolfe, 2007). The assumptions underlying GS4 lead to a quite complex model with asynchronous diffusion processes (Wolfe, 2007). The Poisson random walk fits perfectly to GS4 and could lead to considerably simpler expressions. The mathematical derivation will certainly be more complex than combining two models based on similarity-choice. In any case, we believe that the Poisson random walk with its simplicity and explicative power will prove a very useful tool for the development of cognitive theories.

Relationship to the diffusion model

The Poisson random walk is also closely related to a diffusion model with drift μ and diffusion constant σ by the well-known limit theorem (e.g., Feller, 1968): Let δ be the step size of a single step, p the binomial parameter of the simple random walk ($q = 1 - p$) and r the number of steps per time unit δ . Now, if $\frac{p}{q} \rightarrow 1$, $r \rightarrow \infty$ and $\delta \rightarrow 0$ in a way that

$$(p - q)\delta r \rightarrow \mu \qquad 4pq\delta^2 r \rightarrow \sigma^2$$

then the random walk converges in the limit to a diffusion model with drift μ and diffusion coefficient σ^2 . For the Poisson random walk, this means that the diffusion model is obtained if $\frac{v_A}{v_B} \rightarrow 1$, $C \rightarrow \infty$ and $\delta \rightarrow 0$ in accordance with the expressions for μ and σ^2 above. Taking this limit eliminates one parameter— v_B , or the processing capacity C . In that sense, the diffusion model describes a Poisson random walk with infinite processing capacity. As shown in the previous paragraph, the Poisson random walk can easily be combined with other

theories based on the similarity choice rule; theories that could describe the processing rates of the Poisson random walk in detail. The diffusion model on the other hand is more closely related to signal detection theory (SDT, Tanner & Swets, 1954). Models that are based on and extend SDT form a class with assumptions that differ from the similarity-choice model family. A comprehensive comparison on this distinction can be found in the review by Logan (2004).

Another difference between the diffusion model and the Poisson random walk is the assumption of time inhomogeneous processing. The diffusion model assumes trial-by-trial variation in drift rate, but a constant drift rate within each trial (Ratcliff, 1978). To ease this restriction, the non-decision time is often assumed to vary across trials (Ratcliff & Tuerlinckx, 2002). Effectively, this produces variability in the leading edge of the predicted RT distribution that mimics effects of time inhomogeneous processing. More recently, Smith and Ratcliff (2009) proposed an integrated diffusion model of attention and choice that also includes within trial variation of drift. The model is based on a temporally inhomogeneous diffusion process—that is, a diffusion process with drift that is not constant within a trial but increases (or decreases) over time (Smith, 2000). In that model the drift rate is zero at stimulus onset, grows to an asymptote and decays back to zero once the stimulus has been removed from the display. Obtaining predictions of that model is quite complex but the agreement between theory and data as well as the scope of the integrated diffusion model is remarkable. The Poisson random walk includes time inhomogeneous processing, albeit in a more restricted way than the integrated diffusion model (Smith & Ratcliff, 2009). Processing rates are conceived as the output from sustained (and possibly transient) encoding filters. In all applications of the Poisson random walk presented in this paper, processing rates are zero at stimulus onset and continuously increase to a maximum. This change over time is restricted to proportional hazards, which means that the probabilities of categorization (corresponding to the drift rates in the diffusion model) remain constant throughout the process. As we have demonstrated, inhomogeneous processing with proportional processing rates can easily be

included in the model, but it may be difficult to model time-varying, non-proportional hazards—especially under the additional assumption of trial-by-trial variation in processing rates.

Limitations

Being a theory for response times, the Poisson random walk shares limitations of other RT models while some limitations are avoided. Of the latter, the most interesting improvement is its natural generalization to account for RTs with multiple perceptual categories. Arguably the greatest limitation on the other hand is that the model is mute with respect to the composition of processing rates (“drift rates”) that feed into the random walk. Technically, these are free parameters that need to be estimated from data and thus the number of free parameters inadvertently increases with the number of conditions the model should explain data from. To avoid this, a complementary theory would be needed that describes how the processing rates are composed. We suggested a number of established theories that could achieve this (Logan, 2004). Each of these has its own impressive history in explaining behavioral findings and all are quite general in scope. However, it remains to do the merging and to test how well the resulting theories will explain data. Second, the model so far is a purely cognitive theory of response times. A lot of research has recently focused on linking cognitive theories to neuroscience data and these studies provide converging evidence that sequential sampling models like the diffusion model can also describe decision-making related activity in the brain (Palmeri, Love, & Turner, 2017). The assumption of Poisson processing seems certainly compatible with neural processing (Bundesen, Habekost, & Kyllingsbæk, 2005; Zhang & Bogacz, 2010), however, this association remains speculative at this point. Further studies will be needed to derive quantitative predictions and assess the model on a neural level in order to decide whether the model is useful for cognitive neuroscience as well.

Conclusions

We proposed and tested a response time model for speeded responses based on a random walk and under the assumption of Poisson processing. The resulting model is relatively simple so that an analytic solution is available for two alternatives and numerical methods can be used to evaluate the model for multiple alternatives. The model is derived from the theory of visual attention (Bundesen, 1990). The two models are not only formally related, but also empirically. We provided evidence that the processing rates—central parameters of both models—are empirically indistinguishable. This in turn implies that the notion of categorizations “item x belongs to category i ” is the same both models and that the Poisson random walk paired with TVA assumptions on processing rates provides a general framework for RT-based attention tasks. However, this model combination is not the only one conceivable. In general, all theories based on similarity-choice can be combined with the model presented here with little or no effort. In such combinations the added theory serves as an input to the processing rates of the random walk. The scope and generality of possible combinations are appealing, but the ability to account for empirical data needs yet to be established. The results of the empirical tests presented in this paper are certainly encouraging.

Appendix A

Derivation of the model equations

In this appendix, we present the main ideas behind the Poisson random walk in more detail and provide mathematical derivations its predictions in terms of the cumulative distribution function of first passage times to the lower criterion. It is assumed that the decision is based on Poisson counters that accumulate evidence until a pre-set relative evidence threshold (absorbing barrier) is reached. Assuming two response categories and a relative response rule leads to a simple random walk between two absorbing barriers placed at a and 0 . The quantity of interest is the first passage time to either of these states from z_0 without having reached the opposite criterion before. This problem is known as the gambler's ruin problem (Feller, 1986). The probability function $f_B(n | p, a, b)$ that Player B with initial capital z who repeatedly plays a game to win 1\$ with success probability p (and lose 1\$ with probability $1 - p$) against Player A with initial capital $a - z$ goes bankrupt after exactly n iterations of the game is

$$f_B(n | p, a, z) = \frac{2^n}{a} p^{\frac{n-z}{2}} (1-p)^{\frac{n+z}{2}} \sum_{i=1}^{a-1} \cos^{n-1}\left(\frac{\pi i}{a}\right) \sin\left(\frac{\pi i}{a}\right) \sin\left(\frac{\pi z i}{a}\right) \quad (A1)$$

In cognitive terms, $f_B(n | p, a, z)$ is the probability of reporting perceptual category B after having obtained just n tentative categorizations, with evidence thresholds $k_B = 0$, $k_A = a$, and starting point z . Setting $k_B = k_A = a/2$, (A1) represents the probability to reach the lower criterion k_B (placed at $-a/2$) without having reached either k_A or k_B previously⁸.

The probability p in (A1) is not a free parameter; since the random walk represents the difference in two Poisson counters, it holds that if there occurs a categorization at time t , the probability that it is in Counter A is (e.g., Townsend & Ashby, 1983, p. 273):

⁸ In the common definition (Feller, 1968), a criterion of $a/2$ would only make sense if a is an even integer. As already pointed out in the main text, we accounted for non-integer a by probability mixing. For example, $a = 9$ means that either criterion is 4.5—that is, predictions were obtained by a 50:50 mixture of a model with criteria $[a/2] = 4$ and $[a/2] = 5$. See Appendix C (and Equation (C3)) for details.

$$p = \frac{v_A \exp(-v_A t) \cdot \exp(-v_B t)}{(v_A + v_B) \exp[-(v_A + v_B)t]} = \frac{v_A}{v_A + v_B} \quad (\text{A2})$$

The basic Poisson random walk $X_n = Y_0 + Y_1 + \dots + Y_n$ can therefore be defined as a random walk with increments distributed as

$$P(Y_i = y) = \begin{cases} \frac{v_A}{v_A + v_B}; & y = +1 \\ \frac{v_B}{v_A + v_B}; & y = -1 \end{cases}$$

independent of i , that is, a modified binomial $B\left(1, \frac{v_A}{v_A + v_B}\right)$ distribution where v_A and v_B are the Poisson rates of the two Poisson counters. Inserting (A2) into (A1):

$$f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right) = \frac{2^n}{a} \left(\frac{v_A}{v_A + v_B}\right)^{\frac{n-z}{2}} \left(\frac{v_B}{v_A + v_B}\right)^{\frac{n+z}{2}} S(n \mid a, z) \quad (\text{A3})$$

For convenience, $S(n \mid a, z) := \sum_{i=1}^{2a-1} \cos^{n-1}\left(\frac{\pi i}{a}\right) \sin\left(\frac{\pi i}{a}\right) \sin\left(\frac{\pi z i}{a}\right)$ denotes the finite sum in (A1). The random walk first passage probability (A3) needs to be put into relation to continuous time t to obtain predicted RT distributions of a Poisson random walk. From the Poisson assumption it follows that the interstep time T_n between any two steps in the random walk is exponentially distributed with rate parameter $C = v_A + v_B$ (Bundesen, 1990; Kyllingsbæk et al., 2012; Nosofsky & Palmeri, 1997a) and density function

$$f_{T_n}(t) = C e^{-Ct}$$

for $t > 0$, zero elsewhere. The sum of n exponentially distributed intervals with rate C is Erlang-distributed (gamma distribution with integer shape parameter) with rate C . Hence, $F_T(t)$, the cumulative distribution function (cdf) of waiting time T for n exponentially distributed intervals each with rate C is

$$F_T(t | n, C) = 1 - \exp(-Ct) \sum_{i=0}^{n-1} \frac{(Ct)^i}{i!} \quad (\text{A4})$$

for $t > 0$, zero elsewhere.

Since the number of steps until absorption is unknown it is necessary to marginalize n by summing over all possible values. This yields the cumulative distribution function $F_B(t | v_A, v_B, a, z)$ of first passage time to state $k_B = -a/2$ in continuous time t :

$$F_B(t | v_A, v_B, a, z) = \sum_{n=z}^{\infty} F(t | n, v_A + v_B) f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right) \quad (\text{A5})$$

where $F(t | n, v_A + v_B)$ is an Erlang distribution (A4) with shape parameter n and rate parameter $C = v_A + v_B$, and $f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right)$ given by (A3). Together with the residual latency T_0 , the basic Poisson random walk model contains up to five free parameters $(v_A, v_B, k_A, k_B, T_0)$ that need to be estimated from data.

Trial-by-trial variation in processing rates

The Poisson random walk represents the difference of two Poisson processes. Poisson processing rates are by definition non-negative, so we assume the processing rates to be gamma-distributed:

$$V_A \sim \text{gamma}(\theta_A, \xi)$$

$$V_B \sim \text{gamma}(\theta_B, \xi)$$

The assumption of a common rate ξ is certainly only an approximation but necessary for mathematical tractability. In the remainder of this section, we will derive the Poisson random walk under this assumption.

The assumption of gamma-distributed processing rates leads to a random walk model in which the increments Y_n are distributed as $\text{BB}(1, \theta_A^{(n)}, \theta_B^{(n)})$,

$$P(Y_n = y) = \begin{cases} \frac{\theta_A^{(n)}}{\theta_A^{(n)} + \theta_B^{(n)}}; & y = +1 \\ \frac{\theta_B^{(n)}}{\theta_A^{(n)} + \theta_B^{(n)}}; & y = -1 \end{cases}$$

that is, a beta-binomial distribution with $n = 1$ and parameters $\theta_A^{(n)}, \theta_B^{(n)}$. Hence, the state distribution of the unrestricted process without barriers is $\text{BB}(n, \theta_A, \theta_B)$. The distribution of increments is dependent on n , thus, the process is non-stationary. The initial values are given by θ_A and θ_B , which are increased each time a respective tentative categorization for A or B is made.

Let $G_B(t|\theta_A, \theta_B, \xi, a, z)$ denote the cdf of first-passage time of a Poisson random walk with gamma-distributed processing rates. This cdf is given by the compound distribution

$$G_B(t|\theta_A, \theta_B, \xi, a, z) = \int_0^\infty \int_0^\infty f(v_A|\theta_A, \xi) f(v_B|\theta_B, \xi) \sum_{n=z}^\infty F(t|n, v_A + v_B) f_B\left(n \middle| \frac{v_A}{v_A + v_B}, a, z\right) dv_A dv_B \quad (\text{A6})$$

where $f(v_A|\theta_A, \xi)$ and $f(v_B|\theta_B, \xi)$ represent gamma densities for V_A and V_B , respectively, and the infinite sum is the first-passage time distribution function (cdf) of the Poisson random walk with constant processing rates (A5).

The characteristic dependencies of the random walk first-passage probability function $f_B\left(n \middle| \frac{v_A}{v_A + v_B}, a, z\right)$ and the Erlang distribution $F(t | n, v_A + v_B)$ on the processing rates in

(A6) suggest a change of variables: $U := \frac{v_A}{v_A + v_B}$ and $W := v_A + v_B$. Given the above assumptions of gamma-distributed processing rates, it follows that

$$U \sim \text{beta}(\theta_A, \theta_B)$$

$$W \sim \text{gamma}(\theta_A + \theta_B, \xi)$$

that is, U is beta distributed with parameters θ_A and θ_B ; its density is

$$f(u|\theta_A, \theta_B) = \frac{1}{B(\theta_A, \theta_B)} u^{\theta_A-1} (1-u)^{\theta_B-1}$$

where $B(\theta_A, \theta_B)$ is the beta function with parameters θ_A, θ_B .

Variable W is gamma distributed with parameters $\theta_A + \theta_B$ and ξ ; its density is

$$f(w|\theta_A + \theta_B, \xi) = \frac{w^{\theta_A + \theta_B - 1}}{\Gamma(\theta_A + \theta_B)} \xi^{\theta_A + \theta_B} \cdot \exp(-\xi w).$$

If v_A and v_B are stochastically independent, U and W are also stochastically independent, and their common density is

$$f_{uw}(u, w|\theta_A, \theta_B, \xi) =$$

$$\frac{1}{B(\theta_A, \theta_B)} u^{\theta_A-1} (1-u)^{\theta_B-1} \frac{\xi^{(\theta_A + \theta_B)}}{\Gamma(\theta_A + \theta_B)} w^{(\theta_A + \theta_B - 1)} \exp(-\xi w)$$

By this change of variables, (A6) becomes

$$\begin{aligned}
G_B(t|\theta_A, \theta_B, \xi, a, z) &= \\
&= \int_0^\infty \int_0^1 f_{uw}(u, w|\theta_A, \theta_B, \xi) \sum_{n=z}^\infty F(t|n, w) \cdot f_B(n|u, a, z) du dw
\end{aligned}$$

This considerably simplifies finding a solution to (A6) because the double integral may be rearranged into two separate integrals:

$$\begin{aligned}
G_B(t|\theta_A, \theta_B, \xi, a, z) &= \\
&\sum_{n=z}^\infty \left\{ \int_0^1 f(u|\theta_A, \theta_B) f_B(n|u, a, z) du \cdot \int_0^\infty f(w|\theta_A + \theta_B, \xi) F(t|n, w) dw \right\} \quad (A7)
\end{aligned}$$

The first integral in (A7) is a compound distribution of the random walk first passage time pdf (A3) and the pdf of a beta distribution (Ratcliff & Smith, 2004, p. 367). This product is mathematically tractable as both functions are essentially power functions of u :

$$\begin{aligned}
&\int_0^1 f_B(n|u, a, z) f(u|\theta_A, \theta_B) du = \\
&\frac{2^n}{B(\theta_A, \theta_B)} \frac{S(n|a, z)}{a} \int_0^1 u^{\frac{n-z}{2} + \theta_A - 1} (1-u)^{\frac{n+z}{2} + \theta_B - 1} du
\end{aligned}$$

with $S(n|a, z) = \sum_{i=1}^{a-1} \cos^{n-1}\left(\frac{\pi i}{a}\right) \sin\left(\frac{\pi i}{a}\right) \sin\left(\frac{\pi z i}{a}\right)$. The integral is a beta function:

$$\int_0^1 u^{\frac{n-z}{2} + \theta_A - 1} (1-u)^{\frac{n+z}{2} + \theta_B - 1} du = B\left(\frac{n-z}{2} + \theta_A, \frac{n+z}{2} + \theta_B\right)$$

Hence, the required solution is:

$$\begin{aligned}
& \int_0^1 f(u|\theta_A, \theta_B) f_B(n|u, a, z) du \\
&= \frac{2^n}{a} \frac{B\left(\frac{n-z}{2} + \theta_A, \frac{n+z}{2} + \theta_B\right)}{B(\theta_A, \theta_B)} S(n|a, z)
\end{aligned} \tag{A8}$$

The second integral in (A7) is a compound distribution of an Erlang distribution and a gamma distribution. A compound distribution of this kind has been investigated in terms of its pdf (Dubey, 1970), for the Erlang cdf one must solve

$$\begin{aligned}
& \int_0^\infty f(w|\theta_A + \theta_B, \xi) F(t | n, w) dw = \\
& \int_0^\infty \frac{w^{\theta_A + \theta_B - 1}}{\Gamma(\theta_A + \theta_B)} \xi^{\theta_A + \theta_B} \exp(-\xi w) \left[1 - \exp(-wt) \sum_{i=0}^{n-1} \frac{(wt)^i}{i!} \right] dw,
\end{aligned}$$

the solution of which is

$$\begin{aligned}
& \frac{\xi^{\theta_A + \theta_B}}{\Gamma(\theta_A + \theta_B)} \int_0^\infty w^{\theta_A + \theta_B - 1} \exp(-\xi w) - \exp[-(\xi + t)w] \sum_{i=0}^{n-1} \frac{t^i w^{i + \theta_A + \theta_B - 1}}{i!} dw = \\
& \frac{\xi^{\theta_A + \theta_B}}{\Gamma(\theta_A + \theta_B)} \left[\frac{\Gamma(\theta_A + \theta_B)}{\xi^{\theta_A + \theta_B}} - \sum_{i=0}^{n-1} \frac{t^i \Gamma(i + \theta_A + \theta_B)}{i! (\xi + t)^{i + \theta_A + \theta_B}} \right] = \\
& 1 - \frac{\xi^{\theta_A + \theta_B}}{\Gamma(\theta_A + \theta_B)} \sum_{i=0}^{n-1} \frac{t^i \Gamma(i + \theta_A + \theta_B)}{i! (\xi + t)^{i + \theta_A + \theta_B}}
\end{aligned} \tag{A9}$$

Combining (A8) and (A9) according to (A7) yields the first-passage time cdf of a Poisson random walk with gamma-distributed processing rates:

$$\begin{aligned}
& G_B(t|\theta_A, \theta_B, \xi, a, z) = \\
& \sum_{n=z}^\infty \left[1 - \frac{\xi^{\theta_A + \theta_B}}{\Gamma(\theta_A + \theta_B)} \sum_{i=0}^{n-1} \frac{t^i \Gamma(i + \theta_A + \theta_B)}{i! (\xi + t)^{i + \theta_A + \theta_B}} \right] \frac{2^n}{a} \frac{B\left(\frac{n-z}{2} + \theta_A, \frac{n+z}{2} + \theta_B\right)}{B(\theta_A, \theta_B)} S(n|a, z)
\end{aligned}$$

The associated pdf is:

$$g_B(t|\theta_A, \theta_B, \xi, a, z) =$$

$$\frac{\xi^{\theta_A + \theta_B}}{a \Gamma(\theta_A) \Gamma(\theta_B)} \sum_{n=z}^{\infty} \frac{2^n t^{n-1} \Gamma\left(\theta_A + \frac{n-z}{2}\right) \Gamma\left(\theta_B + \frac{n+z}{2}\right)}{(n-1)! (\xi + t)^{n+\theta_A+\theta_B}} S(n|a, z)$$

The pdf may be used for maximum likelihood estimation (Van Zandt, 2000; Myung 2003) or determining the hazard rate. Under the assumption of a common shape parameter ξ for the distribution of processing rates the obtained solution has only a single parameter more than the Poisson random walk model with constant processing rates v_A and v_B . The assumption of gamma-distributed processing rates makes the process non-stationary, since its increments become dependent on n —the number of steps in the random walk.

Time inhomogeneous processing

In the above derivations, it was assumed that processing rates are constant *within* a trial (constant hazard rates). This restriction can be relaxed to obtain a time inhomogeneous version of the model where the processing rates may change over time. Let

$$v(t) = v(x, i, t)$$

be the time-dependent processing rate for the categorization “stimulus x belongs to category i ”. In the following the indices x and i will be omitted for the sake of clarity and because the rates of all categorizations are assumed to follow the same time course. The following derivations are not limited to the Poisson random walk; rather they are adapted from the more general concept of time-dependent Poisson processes with event rate $\lambda(t)$ (Smith & Van Zandt, 2000). Accordingly, let

$$\lambda(t) := v(t)$$

be the event rate of a time dependent Poisson counter, and

$$\Lambda(t) := \int_0^t v(s) ds$$

its integrated event rate (Smith & Van Zandt, 2000).

The encoding filter output exhibits phasic or tonic response profiles—that is, activation from transient and sustained encoding channels (Smith, 1998; Smith & Van Zandt, 2000). For the Poisson random walk model, it turned out that sustained encoding filters were sufficient to model RT distributions. We modelled the response profile of sustained filters with a gamma distribution function—that is, the temporal profile of the output of sustained encoding channels is represented by the gamma cdf (Smith & Van Zandt, 2000, their Eq. 20):

$$\lambda_s(t)/v_s = 1 - \exp(-\mu_s t) \sum_{j=0}^{n_s-1} \frac{(\mu_s t)^j}{j!}$$

where v_s is the amplitude of sustained encoding filters and μ_s and n_s define their temporal characteristics. In principal, each perceptual category can have different temporal characteristics, hence, different μ_s and n_s . To keep things simple, we assumed that these parameters were identical for all perceptual categories. Under this assumption, the event rates are temporally inhomogeneous, but their ratio remains constant over time (proportional hazards, Bundesen, 1993). The integrated event rate is

$$\frac{\Lambda_s(t)}{v_s} = \bar{\Lambda}_s(t) = \int_0^t \left[1 - \exp(-\mu_s s) \sum_{j=0}^{n_s-1} \frac{(\mu_s s)^j}{j!} \right] ds = t - \sum_{j=0}^{n_s-1} \frac{\gamma(\mu_s t, j+1)}{\mu_s} \quad (\text{A10})$$

with $\gamma(x, k)$ representing the regularized lower incomplete gamma function. In the full model, the filter amplitude v_s is assumed to exhibit trial-by-trial variability, so it is convenient to define a normalized integrated event rate, $\bar{\Lambda}_s(t)$. In all fits of the model, we fixed $n_s = 5$ (see Smith & Van Zandt, 2000, p. 312) and estimated μ_s from data.

Inserting (A10) into the expressions for the Erlang distribution (A4) for time until absorption given there are n steps and overall processing capacity C yields

$$F[\bar{\Lambda}_S(t) | n, C] = 1 - \exp[-C\bar{\Lambda}_S(t)] \sum_{i=0}^{n-1} \frac{C^i [\bar{\Lambda}_S(t)]^i}{i!}.$$

The associated pdf is

$$f[\bar{\Lambda}_S(t) | n, C] = \frac{C^n [\bar{\Lambda}_S(t)]^{n-1} \bar{\Lambda}_S(t) \exp[-C\bar{\Lambda}_S(t)]}{(n-1)!}$$

(see Smith & Van Zandt, 2000, their Eqs. 3 and 4). The temporal profile of encoding channels is assumed to be the same for all counters. Under this simplifying assumption, the first passage time probability (A1) does not change, and the first passage time cdf with time inhomogeneous processing but *without* trial-by-trial variation in processing rates is

$$F_B(\bar{\Lambda}_S(t) | v_A, v_B, a, z) = \sum_{n=z}^{\infty} \left\{ 1 - \exp[-(v_A + v_B)\bar{\Lambda}_S(t)] \sum_{i=0}^{n-1} \frac{(v_A + v_B)^i [\bar{\Lambda}_S(t)]^i}{i!} \right\} \cdot f_B\left(n \middle| \frac{v_A}{v_A + v_B}, a, z\right) \quad (\text{A11})$$

The full model

What remains to be shown is the derivation of the full model—that is, the Poisson random walk model with trial-by-trial variation of processing rates and evidence thresholds together with the assumption of time inhomogeneous processing. With time-dependent processing rates, the trial-by-trial variation assumption in processing rates corresponds to the assumption that the amplitude of the sustained encoding filters is gamma-distributed:

$$V_S \sim \text{gamma}(\theta_S, \xi)$$

We start the derivation of the full model with Equation (A11) and note that the necessary calculations are identical to those in the time homogeneous model: Only the Erlang-distribution changes through the assumption of time-dependent processing rates; the first-passage probability of the random walk is unaffected because the $\bar{\lambda}_S$ cancel out (see Equation(A2)). With the definition of a normalized integrated event rate, $\bar{\lambda}_S$, it is immediately seen that the Erlang distribution with gamma-distributed processing rates is obtained by inserting $\bar{\lambda}_S(t)$ for t (see derivation of (A9)). Therefore,

$$\begin{aligned}
 G_B[\bar{\lambda}_S(t)|\theta_A, \theta_B, \xi, a, z] = & \\
 & \int_0^\infty \int_0^\infty f_{uv}(u, v|\theta_A, \theta_B, \xi) \sum_{n=z}^\infty F[\bar{\lambda}_S(t)|n, x+y] \cdot f_B\left(n \middle| \frac{v_A}{v_A+v_B}, a, z\right) du dv = \\
 & \frac{1}{a B(\theta_A, \theta_B)} \sum_{n=z}^\infty \left[1 - \frac{\xi^{\theta_A+\theta_B}}{\Gamma(\theta_A+\theta_B)} \sum_{i=0}^{n-1} \frac{\Gamma(i+\theta_A+\theta_B)}{i!} \frac{[\bar{\lambda}_S(t)]^i}{[\xi + \bar{\lambda}_S(t)]^{i+\theta_A+\theta_B}} \right] \times \\
 & 2^n B\left(\frac{n-z}{2} + \theta_A, \frac{n+z}{2} + \theta_B\right) S(n|a, z)
 \end{aligned} \tag{A12}$$

This is the first-passage time cdf of the model with gamma-distributed processing rates and time inhomogeneous processing. As before, $S(n|a, z)$ represents the finite sum in the first passage probability (A1) and $\bar{\lambda}_S(t)$ is the normalized integrated event rate given by (A10).

The full model presented in the main text is obtained using (A12) and including trial-by-trial variation in the starting point z . Since the increments are beta-binomial distributed, the starting point z is assumed to be beta-binomial distributed as well:

$$Z \sim \text{BB}(n_z, z_A, z_B)$$

with probability function

$$P(Z = z | n_z, z_A, z_B) = \binom{n_z}{z-1} \frac{B(z_A + z - 1, z_B + n_z - z + 1)}{B(z_A, z_B)}$$

for $z \geq 1$ and $z \leq n_z$, zero elsewhere. Parameters z_A and z_B are free parameters that need to be estimated from data. Parameter n_z was fixed ($n_z = a - 1$) so that $P(z)$ was non-zero only for $z = (1, 2, \dots, a - 1)$ to avoid degenerations of the process (absorption before stimulus onset).

Let $H_B[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z_A, z_B]$ represent the first-passage time cdf of the full model with beta-binomial distributed starting point (z_A, z_B) , gamma distributed processing rates $(\theta_A, \theta_B, \xi)$ and time inhomogeneous processing (μ_S) . The starting point is an integer valued variable in the Poisson random walk; hence, combining (A12) with a beta-binomial starting point distribution $P(z | a - 1, z_A, z_B)$ yields

$$H_B[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z_A, z_B] = \sum_z G_B[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z] \times P(z | a - 1, z_A, z_B)$$

Thus, starting point variability is included by summing over all possible starting point values the probabilities $P(z | a - 1, z_A, z_B)$ of which follow a beta-binomial distribution. For the alternative categorization, the cumulative distribution function of first-passage times to the upper criterion, $H_A[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z_A, z_B]$, is obtained by the parameter substitution $H_A[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z_A, z_B] = H_B[\bar{\Lambda}_S(t) | \theta_B, \theta_A, \xi, a, z_B, z_A]$.

In the models with variable non-decision time T_0 , we assumed T_0 to be uniformly distributed with width s_t (Ratcliff & Tuerlinckx, 2002):

$$J_B[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z_A, z_B, s_t] = \int_{T_0 - s_t/2}^{T_0 + s_t/2} H_B[\bar{\Lambda}_S(t - s) | \theta_A, \theta_B, \xi, a, z_A, z_B] \cdot \frac{1}{s_t} ds$$

We could not find an analytical solution for $J_B[\bar{\Lambda}_S(t) | \theta_A, \theta_B, \xi, a, z_A, z_B, s_t]$, so we used numerical methods to evaluate the model with variable non-decision time.

Model of pure accuracy data

The Poisson counter model for pure accuracy data accounts for time-controlled processing and does not entail decision criteria. Instead, evidence is sampled until time t when processing is interrupted by mask onset. The probability P_i that a total of n tentative categorizations “item x belongs to category i ” are made up to time t follows a Poisson distribution:

$$P_i[n | v(x, i)t] = \frac{[v(x, i)t]^n}{n!} \exp[-v(x, i)t]$$

The probability of reporting category i when presented with stimulus x is a sum of three probabilities:

$$P(x, i) = P_1(x, i) + P_2(x, i) + P_3(x, i)$$

With two perceptual categories, say i and j , the probability $P_1(x, i)$ corresponds to the probability of counter i having a higher count than counter j , $P_2(x, i)$ to the probability that both counters have equal (non-zero) counts and the participant guesses category i , and $P_3(x, i)$ to the probability of both counters being zero and guessing category i at random. The general formulae for those probabilities with any number of report categories are given in Kyllingsbæk et al. (2012); we will briefly state them in the simplified form for two perceptual categories.

The probability $P_1(x, i)$ for counter i being higher than counter j is

$$P_1(x, i) = \sum_{n=1}^{\infty} \frac{[v(x, i)t]^n}{n!} \exp[-v(x, i)t] \times \sum_{m=0}^{n-1} \frac{[v(x, j)t]^m}{m!} \exp[-v(x, j)t],$$

the probability $P_2(x, i)$ for both counters being equal and non-zero *and* randomly selecting category i is

$$P_2(x, i) = .5 \times \sum_{n=1}^{\infty} \frac{[v(x, i)t]^n}{n!} \exp[-v(x, i)t] \times \frac{[v(x, j)t]^n}{n!} \exp[-v(x, j)t],$$

and the probability $P_3(x, i)$ for both counters being zero and blindly guessing category i is

$$P_3(x, i) = p_i \times \exp\{-[v(x, i) + v(x, j)]t\}$$

where p_i is the probability of blindly guessing category i —that is, guessing i in absence of any information about the stimulus x .

Under the assumption of gamma-distributed processing rates, the Poisson distribution of the number of counts in counter i becomes a negative binomial distribution:

$$\begin{aligned} P_i[n | t, \theta(x, i), \xi] &= \int_0^{\infty} f[v | v_{\theta}(x, i), \xi] P_i[n | t, v] dv \\ &= \frac{\Gamma[n + \theta(x, i)]}{n! \Gamma[\theta(x, i)]} \frac{t^n \xi^{\theta(x, i)}}{(t + \xi)^{n + \theta(x, i)}} \end{aligned}$$

A time inhomogeneous Poisson counter model is obtained by exchanging the rate parameter vt of the Poisson distribution with the integrated event rate of the time-dependent event rate $v(x, i, t)$:

$$P_i[n | t, v(x, i, t)] = \frac{\left[\int_0^t v(x, i, s) ds \right]^n}{n!} \exp \left[- \int_0^t v(x, i, s) ds \right] = \frac{[\Lambda(t)]^n}{n!} \exp[-\Lambda(t)]$$

This is a temporally generalized Poisson distribution (Smith & Van Zandt, 2000). As in the RT model, we assume an encoding component $\lambda_S(t)$ that has a sustained response profile. Additionally, an encoding component with a transient profile $\lambda_T(t)$ was included; their temporal profile was modeled with a gamma pdf (Smith & Van Zandt, 2000):

$$\lambda_T(t)/v_T = \frac{(\mu_T t)^{n_T-1}}{(n_T-1)!} \mu_T \exp(-\mu_T t)$$

The amplitude of transient filters is v_T ; in contrast to the sustained channels, we conceive the transient filters to contribute unspecific stimulus information (Christensen et al., 2018), that is, the filter amplitude v_T was identical for both counters, violating the proportional hazards assumption. The integral of event rate $\lambda_T(t)$ is simply the gamma cdf:

$$\begin{aligned} \frac{\Lambda_T(t)}{v_T} &= \bar{\Lambda}_T(t) = \int_0^t \frac{(\mu_T s)^{n_T-1}}{(n_T-1)!} \mu_T \exp(-\mu_T s) ds \\ &= 1 - \exp(-\mu_T t) \sum_{i=0}^{n_T-1} \frac{(\mu_T t)^i}{i!} \end{aligned} \quad (\text{A13})$$

The combined contribution of both transient and sustained channels is conceived as the sum

$$\bar{\Lambda}(t) = \gamma \bar{\Lambda}_T(t) + \bar{\Lambda}_S(t) \quad (\text{A14})$$

with relative transient channel gain $\gamma = \frac{v_T}{v_S}$. Gain γ represents the amplitude of a transient channels filter relative to a sustained channels filter. In the model with trial-by-trial variation in V_S and V_T , we set $V_T := \gamma_\theta V_S$, that is, we assumed the ratio $\gamma_\theta = \frac{\theta_T}{\theta_S}$ to be constant. In other words, V_S and V_T co-varied across trials but were stochastically independent for the different perceptual categories.

Appendix B

The Poisson random walk for n perceptual categories

Since the superposition principle of Poisson processes is not restricted to just two counters, it is possible to formulate an extension to multiple perceptual categories. As in the main text, we illustrate the generalization using the special case of three categories but adding more categories is straightforward. With three perceptual categories for x , evidence for each is

stored in a separate Poisson counter i with rate v_i . The time between any two tentative categorizations is exponentially distributed with rate parameter $C = \sum v_i$ so that the time to obtain n events is Erlang-distributed with rate C (cf. (A4)). This part of the Poisson random walk is easily generalized to any number of alternatives; the problem is to find a general expression for the first-passage time probability function of the random walk (A1).

As in the binary categorization model we assume that evidence for one category is evidence against all other categories. In principle, one could formulate such a model with three competing Poisson counters that are increased if a respective tentative categorization for x is made and decreased if any other categorization is made. Without a further modification, however, this does not lead to a well-founded model: inhibition will outweigh activation and it is not generally the case that any of the evidence criteria is ever reached at all. Therefore, competitive networks of this type are usually supplied with an auxiliary assumption, namely that inhibition can never lead to negative activation. Thus, the state $x = 0$ in each counting process could be considered a reflecting barrier (e.g., Cox & Miller, 1965) meaning that from there the process can only reach state $x = 1$ or remain in $x = 0$. The states that can possibly be reached are thus confined by this reflecting barrier and the evidence criterion (absorbing barrier) of each counter. Each time a tentative categorization is made, the probability for each of the three possible categorizations ($i = 1, 2, 3$) is

$$p_i = \frac{v_i}{\sum_j v_j} \quad (\text{B1})$$

As in the binary categorization model, all processes are connected to each other at state $x = 0$, so is possible to map the state space of all processes onto a line—just as in the two-choice case. With three alternatives, the state space is simply divided into three non-contiguous compartments. From the neutral state, the process jumps into one of these compartments, depending on the tentative categorization that is made next. Since the Poisson

random walk model has a discrete state space, these jumps do not differ fundamentally from the other transitions. A schematic representation of a competing network of three Poisson counters is given in the main text. A suitable transition matrix \mathbf{T} of the process could be defined as

$$\mathbf{T} = \begin{bmatrix} 0 & p_A & 0 & 0 & \cdots & 0 & p_B & 0 & 0 & \cdots & 0 & p_C & 0 & 0 & \cdots & 0 \\ q_A & 0 & p_A & & & & & & & & & & & & & 0 \\ 0 & q_A & 0 & p_A & & & & & & & & & & & & 0 \\ \vdots & & \ddots & \ddots & \ddots & & & & & & & & & & & \vdots \\ 0 & \cdots & & q_A & 0 & p_A & & & & & & & & & & 0 \\ 0 & \cdots & & & 0 & 1 & 0 & & & & & & & & \cdots & 0 \\ q_B & 0 & \cdots & & 0 & 0 & p_B & & & & & & & & \cdots & 0 \\ 0 & \cdots & & & & q_B & 0 & p_B & & & & & & & \cdots & 0 \\ \vdots & & & & & & \ddots & \ddots & \ddots & & & & & & & \vdots \\ 0 & \cdots & & & & & & q_B & 0 & p_B & & & & & & 0 \\ 0 & \cdots & & & & & & & 0 & 1 & 0 & & & & \cdots & 0 \\ q_C & 0 & \cdots & & & & & & & 0 & 0 & p_C & & & \cdots & 0 \\ 0 & \cdots & & & & & & & & q_C & 0 & p_C & & & \cdots & 0 \\ \vdots & & & & & & & & & & \ddots & \ddots & \ddots & & & \vdots \\ 0 & & & & & & & & & & & q_C & 0 & p_C & & \\ 0 & \cdots & & & & & & & & & \cdots & 0 & 0 & 1 \end{bmatrix} \quad (\text{B2})$$

with transition probabilities p_i given by (B1) and $q_i = 1 - p_i$. The first row contains transition probabilities from the neutral state, the first column transition probabilities to the neutral state. The three rows with unit probability contain transition probabilities from the absorbing states. The second row and the two rows below the absorbing states contain transition probabilities from the first state in each of the three counters. The state probabilities $P(X=x)$ of this process after just n steps are calculated as

$$P(X=x|n, \mathbf{v}, \mathbf{k}) = \mathbf{z} \cdot \mathbf{T}^n \quad (\text{B3})$$

with transition matrix (B2) and starting point probability vector z . For the fits presented in this paper we used $z = (1, 0, \dots, 0)$, that is, the process starts in the neutral state. Other starting point distributions, for example, based on the multinomial distribution, are conceivable.

Let $P_i(n)$ represent the probability to reach criterion k_i after exactly n steps for the first time without having previously reached any other criterion k_j (first passage time). The required first passage probabilities for the three alternatives are

$$P_i(n|\mathbf{v}, \mathbf{k}) = dP(X = k_i|n, \mathbf{v}, \mathbf{k})/dn$$

The expressions on the right can be obtained using (B3). Derivation w.r.t n is necessary since we are interested in the probability to reach the absorbing barriers after *exactly* n steps.

Finally, the RT model predictions for the three alternatives are

$$F_i(t|\mathbf{v}, \mathbf{k}) = \sum_{n=k_i}^{\infty} F(t|n, C) \cdot P_i(n|\mathbf{v}, \mathbf{k}) \quad (\text{B4})$$

where $F(t|n, C)$ is the cdf of an Erlang distribution (A4) with shape parameter n and rate parameter C . Some basic properties of the model are presented in the main text.

The Poisson random walk for multiple categorizations can also be extended in the same ways as the binary Poisson random walk. From (B4) it is evident that the model is readily generalized to time inhomogeneous processing with constant hazard ratios, since this assumption only changes the Erlang distribution $F(t|n, C)$. Starting point variation could be included by providing a suitable distribution for z in (B3). Trial-by-trial variation of processing rates can be included by specifying the appropriate transition probabilities in (B2). As stated in the main text, if the processing rates are assumed to be gamma distributed, the increments of the model correspond to a Polyá urn model. In the binary categorization model the (non-stationary) transition probabilities $p_i^{(n)}$ for the transition matrix follow a beta binomial distribution. In the general case, the transition probabilities follow a Dirichlet-

multinomial distribution. This process is non-stationary; the $p_i^{(n)}$ depend on the number of steps n . It is therefore necessary to update the transition matrix after each step.

Appendix C

In this appendix, we provide details on the fitting procedure used to estimate model parameters of the fits presented in the main text. The purpose of this appendix is to provide sufficient details on the fitting procedure to replicate the model fits presented in the main text.

Model specification

Experiment 1 contained six different stimuli (x) that were to be categorized as belonging to either of two perceptual categories ($i = A, B$). Following the usual specification in TVA, we devised the processing rates as

$$v(x, i) = \eta(x, i) \beta_i$$

As stated in the main text, for the model with trial-by-trial variability, we did not specify the exact distributions of perceptual evidence and bias. Instead, we modelled the effect of their combined variation on $v(x, i)$ by a gamma distribution with mean

$$\bar{v}(x, i) = \frac{\eta_\theta(x, i) \beta_\theta(i)}{\xi}$$

and scale $1/\xi$. It turned out to be a reasonable approximation to assume that the perceptual evidence parameter ($\eta_\theta(x, i)$) for a correct categorization was the same for both categories, so $\eta_\theta(x, i)$ varied only with confusability. The same restriction was applied to the perceptual evidence parameter of incorrect categorizations. For the bias parameter, we fixed $\beta_\theta(2) = 1$. Together, these seven parameters yielded the 12 processing rates for the model to predict 12 RT distributions (i.e., correct and incorrect distributions to all six stimuli).

In the pure accuracy task we used standard maximum likelihood estimation—that is, we estimated the parameter values by maximizing

$$LL = N \sum_{i,j,k} p_{ijk} \ln(\pi_{ijk})$$

with $N = 60$ (replications in each condition), and the observed p_{ijk} and predicted π_{ijk} probabilities of report in orientation condition k ($k = 1, \dots, 6$), exposure duration j and report i ($i = A, B, 0$).

For the RT fits, we used the quantile maximum likelihood estimation procedure to fit the model to empirical cumulative distribution functions. As stated in the main text, we estimated the 10%, 30%, 50% (median), 70% and 90% quantile from data and fitted the model to these quantiles:

$$\begin{aligned} \text{QML} = \sum_i n_{0.1}^{(i)} \log[F_i(x_{0.1})] + (n_{0.3}^{(i)} - n_{0.1}^{(i)}) \log[F_i(x_{0.3}) - F_i(x_{0.1})] + \dots \\ + (n^{(i)} - n_{0.9}^{(i)}) \log[P_i - F_i(x_{0.9})] \end{aligned} \quad (\text{C2})$$

where i represents the two report categories ($i = A, B$).

For the RT fits of Experiment 1, we maximized (C2) using simplex optimization. For the common fit of RT and pure accuracy data we maximized the sum of (C1) and (C2). The simplex optimization was restarted several times (1,500 model evaluations per iteration) to avoid finding a local minimum. The whole procedure was ended when the fit improved only marginally between two iterations; we chose $< 1e-3$ as exit criterion. The same procedure was used for the RT fits of Experiment 2. The quantile maximum likelihood is readily applicable to any number of alternatives. We maximized (C2) which was a sum across four categories ($i = 1, 2, 3, 4$) instead of two.

Evaluation of the model function

As stated in the main text and Appendix A, we used a probability mixing approach to fit integer-valued parameters, for example, the criterion separation a in the full model:

$$\begin{aligned}
H_B[\bar{\Lambda}_S(t)|\theta_A, \theta_B, \xi, a, z_A, z_B] &= \\
&= H_B[\bar{\Lambda}_S(t)|\theta_A, \theta_B, \xi, \lceil a \rceil, z_A, z_B]P(\lceil a \rceil) + H_B[\bar{\Lambda}_S(t)|\theta_A, \theta_B, \xi, \lfloor a \rfloor, z_A, z_B]P(\lfloor a \rfloor). \quad (C3)
\end{aligned}$$

That is, we obtained predictions for a non-integer value a by calculating predictions of a model in which a assumed the ceiling value ($\lceil a \rceil$) and predictions in which it assumed the floor value ($\lfloor a \rfloor$). The probability weights are $P(\lceil a \rceil) = \lceil a \rceil - a$ and $P(\lfloor a \rfloor) = 1 - P(\lceil a \rceil)$.

The model functions contain an infinite series (e.g., Equation (A5)). For the model fits, the infinite series had to be truncated to obtain predictions:

$$F_B(t | v_A, v_B, a, z) = F_B^{(N)}(t | v_A, v_B, a, z) + \varepsilon$$

with

$$F_B^{(N)}(t | v_A, v_B, a, z) = \sum_{n=z}^N F(t | n, v_A + v_B) f_B\left(n \mid \frac{v_A}{v_A + v_B}, a, z\right)$$

To control truncation error ε , the number of terms N must be chosen so that

$$\left| F_B(t | v_A, v_B, a, z) - F_B^{(N)}(t | v_A, v_B, a, z) \right| < \varepsilon$$

This was achieved by obtaining N from the inverse of a Poisson cdf with rate Ct :

$$N = F^{-1}(1 - \varepsilon | Ct) \quad (C4)$$

where C is the Poisson random walk parameter of visual processing capacity and t is the maximum of all time values required in an evaluation. The same approach was used for the time inhomogeneous models, for which Ct in (C4) was exchanged by the integrated event rate $\Lambda(t)$. For the models with gamma-distributed processing rate, we also used this approach, but calculated N from the inverse cdf of a negative binomial distribution. This approach of truncating the infinite series not only guarantees a pre-set numerical precision, but also

considerably speeds up calculations since all predictions can be obtained in a single step if N is known in advance (Blurton, Kesselmeier, & Gondan, 2014). We required a numerical precision of $\varepsilon = 1\text{e-}5$ in all model fits.

Appendix D

Comparison with the diffusion model of response time

The diffusion model (Ratcliff, 1978; Ratcliff, 1981, Ratcliff & Rouder, 1998) is a continuous time, continuous state space equivalent to the simple random walk. According to the diffusion model, decisions are based on diffusion processes reflecting moment-to-moment fluctuations that develop over time until one of two response criteria is reached. In the usual parametrization of the diffusion model, these criteria are located at a and 0 , and the diffusion process starts at z . The diffusion process is a stochastic process (Wiener process) with normally distributed $N(\mu, \sigma^2)$ increments. Parameter μ is the deterministic tendency of the process, called drift, whereas σ^2 represents the decision noise of the process. The drift rate is specific to a stimulus; the higher the drift the faster and more accurate are the predicted responses. The diffusion constant σ^2 is merely a scaling factor (usually fixed at $\sigma^2 = 0.01$). As in the Poisson random walk, drift rates and starting point are assumed to vary from trial to trial: drift rates are assumed to be distributed as $N(\nu, \eta^2)$ and the starting point as uniformly distributed around z with width s_z .

We fitted both the diffusion model (Ratcliff, 1978; Ratcliff & Rouder, 1998) and the Poisson random walk model to RT distributions obtained in a visual categorization experiment with speed-accuracy instructions for a detailed comparison. Ratcliff et al. (2001) report the effect of aging on response times in a visual categorization task. The study is highly suitable for our purpose, as the stimuli used in that task were clearly mutually confusable. Two groups of participants (young and older participants) were given the task to categorize a

visual display containing a varying number of dots as “high” or “low”, depending on whether the display contained a high or low number of dots (for more experimental details, see Ratcliff et al., 2001). Responses of varying confusability conditions were pooled into four conditions and the obtained RT distributions were averaged across participants.

For the sake of comparability, we used the uniform distribution to include starting point variation (width s_z) in both models. A symmetric ($z = a/2$) diffusion model was fitted to data in that study, the same restriction ($k_A = k_B = a/2$) was made in the Poisson random walk. The diffusion model fit did not include between-trial variation in non-decision time, so we specified the Poisson random walk model with a constant T_0 and time homogeneous processing. To account for varying confusability of the four combined RT conditions only the drift/processing rate was allowed to vary across conditions. The speed-accuracy instructions were assumed to affect only criterion separation a , but not drift rate distributions or non-decision time. We specified the Poisson random walk model to exactly match all these assumptions. Variation in drift rate (diffusion model) was modelled with a normal distribution ($N(v_i, \eta^2)$), variation in processing rates (Poisson random walk) with a gamma distribution ($\text{gamma}(\theta_i, \xi)$). The full Poisson random walk would have too many parameters to obtain a meaningful comparison. Therefore, we restricted visual processing capacity $C = \sum v_i$ to be invariant across stimulus conditions. Under the assumption of gamma distributed processing rates, this corresponds to the distribution parameter θ_C being equal in all conditions. In total, the Poisson random walk had one parameter (θ_C) more than the diffusion model. Parameter differences between age groups were tested for significance by nested model tests.

Table D-1.

Parameter estimates of the Poisson random walk and the diffusion model for the data reported in Ratcliff, Thapar, & McKoon (2001; Exp. 2).

Poisson random walk model parameter	Estimate		Diffusion model parameter	Estimate (Ratcliff et al. results in parentheses)	
	young partici- pants	older partici- pants		young partici- pants	older partici- pants
$\bar{\nu}_1$ (in Hz)	53.35	52.48	ν_1	0.355 (0.330)	0.376 (0.330)
$\bar{\nu}_2$ (in Hz)	47.58	45.53	ν_2	0.220 (0.205)	0.208 (0.206)
$\bar{\nu}_3$ (in Hz)	43.44	41.94	ν_3	0.123 (0.108)	0.120 (0.136)
$\bar{\nu}_4$ (in Hz)	40.06	38.92	ν_4	0.049 (0.043)	0.046 (0.036)
\bar{C} (in Hz)	76.31	74.07	—		
$1/\xi$	1.83	1.79	η	0.139 (0.104)	0.143 (0.108)
k <i>speed</i>	3.78*	4.42*	a <i>speed</i>	0.091 (0.091)	0.109 (0.107)
<i>accuracy</i>	6.5*	8.0*	<i>accuracy</i>	0.155 (0.148)	0.193 (0.181)
s_z	1.72	3.26	s_z	0.039 (0.040)	0.057 (0.045)
T_0 (in s)	0.292*	0.342*	T_{er} (in s)	0.289 (0.285)	0.342 (0.323)

Note—*Poisson random walk model*: Parameters $\bar{\nu}$ and \bar{C} denote mean processing rates and mean processing capacity, respectively (obtained from gamma distribution with scale $1/\xi$). Criterion k represents the evidence threshold for both alternatives ($k = k_1 = k_2$). Parameter s_z is the width of the starting point distribution (uniform distribution around $z = k$). Parameters k and s_z are discrete; real valued parameter estimates denote probability mixing of model predictions using floor and ceiling functions of those values. Parameter T_0 is (constant) residual/motor time. *Diffusion model*: Parameters ν represent mean drift rates and η the trial-by-trial standard deviation of drift (normal distribution). Parameter a is criterion separation, s_z is the width of the starting point distribution (uniform distribution around $z = a/2$). Parameter T_{er} is (constant) non-decision time. Diffusion constant (scaling parameter) was set to $\sigma = 0.1$. Parameters marked with an asterisk (*) are significantly ($p < .05$) different between age groups.

For the diffusion model, the re-analysis yielded estimates close to those reported in the original publication (Ratcliff, et al., 2001). The small differences are likely due to the fact that the number of responses in each of the four response conditions was unknown (these are not reported), so we assumed them to be equal. For this reason, we report the re-fitted diffusion model to compare both models under this assumption. Overall, the diffusion model fit is very similar to the fit originally reported: predictions of the diffusion model agree with the data from correct and error responses in all conditions except for the highest (i.e., 90%) quantiles of some error RT distributions (Figure D-1, A–D, see also Ratcliff et al., 2001).

The cognitive interpretation of both models is identical: Both the older and the younger participants changed their evidence threshold in accordance with the instruction. In the condition in which response speed was stressed, the group of young (older) participants reported a perceptual category after obtaining $k = 3.8$ (4.4) categorizations for the correct perceptual category, whereas $k = 6.5$ (8.0) categorizations were obtained when response accuracy was stressed. Both the evidence thresholds for accurate response condition (3.78 vs. 4.42, $\chi^2(1) = 37.3$, $p < .001$) and speeded response condition (6.5 vs 8.0, $\chi^2(1) = 80.8$, $p < .001$) are significantly different between age groups. Visual processing capacity and all processing rates were slightly higher for the group of young participants, but none of these differences is significant. The residual latency is significantly lower for younger participants (0.292 vs. 0.342 s, $\chi^2(1) = 144.8$, $p < .001$). In sum, the faster responses observed in the group of young participants compared to that of older participants is a combined effect of higher response caution and higher residual latency (“motor time”) in the group of older participants, whereas the speed of perceptual processing was largely similar. Altogether, these results reflect those of the diffusion model (Ratcliff et al., 2001).

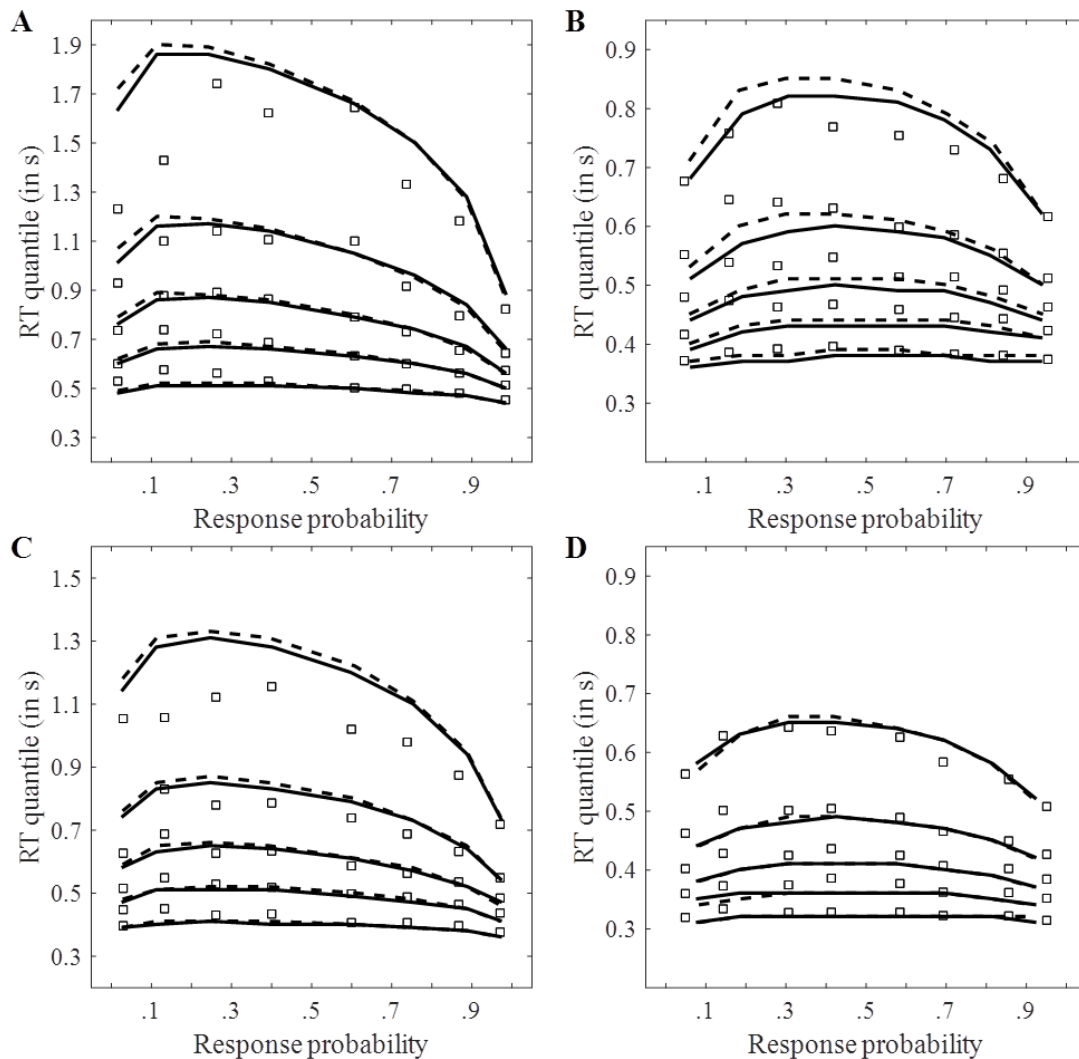


Figure D-1. Data from Ratcliff et al. (2001, their Exp. 2) and predictions of the diffusion model (re-fit, unmarked points connected with dashed lines) and the Poisson random walk model (unmarked points connected with solid lines). In each panel, correct and error RT distributions in the four stimulus conditions are represented by their 10%, 30%, 50% (median), 70%, and 90% quantile (from bottom to top). The four left distributions in each panel are error RT distributions; the four distributions on the right are correct RT distributions. Upper panels (A, B) are data and model predictions for the group of older participants, lower panels (C, D) for the group of young participants. Left panels are model predictions and data obtained under accuracy instruction, right panels under speed instruction. In both models, the (mean) drift/processing rate varied across the four confusability conditions and the criterion varied across speed and accuracy instruction conditions. Model parameters estimates for both models are summarized in Table D-1. Table D-1. Adapted from “The effects of aging on reaction time in a signal detection task” by R. Ratcliff, A. Thapar, and G. McKoon, 2001, *Psychology and Aging*, 16, p. 331. Copyright 2001 by American Psychological Association. Adapted with permission.

The Poisson random walk provides an equally good account of the data; its predictions closely match those of the diffusion model and the only notable mismatch between theory and data occurred in the tail of some error distributions (90% quantiles) which the diffusion model also failed to predict. Quantitatively, the Poisson random walk model fit is slightly better than the diffusion model, with the Poisson model having one additional free parameter (processing capacity θ_C).

Appendix E

Response times in absolute identification

Multi-alternative models such as the Poisson random walk model are also suited to model RT distributions in absolute identification. We fitted the model to a dataset obtained in length identification (Lacouture & Marley, 2004) with then perceptual categories. In Experiment 2 of that study (Lacouture & Marley, 2004) a single observer was instructed to identify the length of vertical bars. There were stimuli of ten different lengths, thus, ten report categories.

Responses were given verbally or manually; as these did not differ fundamentally, we pooled data from both response modalities (see Brown & Heathcote, 2008, for a similar approach). A full model of the dataset would contain a hundred processing rates. This number could be substantially reduced by using Equation (1) to specify mean processing rates and employing several simplifying assumptions. Firstly, we restricted the perceptual evidence $\eta(x, C)$ for all correct categorizations to be equal. Perceptual bias was different across perceptual categories, that is, we fitted a model with ten bias parameters ($\beta_1, \beta_2, \dots, \beta_{10}$), nine of which were free parameters ($\beta_{10} = 1$). Secondly, we assumed the visual processing capacity C to be equal across all x so that the sum of processing rates for all erroneous categorizations were obtained as $C - v(x, C)$, that is, the remainder of processing capacity after subtracting the processing rate for the correct categorization (C). Thirdly, note that the confusion matrix is sparse,

practically only errors of ± 1 unit were observed—that is, errors occurred mostly as confusions between the correct category and either the next shorter or the next longer perceptual category. Other errors were very rare and neglected for this fit. We used this to restrict the perceptual evidence to be equal for the two adjacent incorrect categories. For stimuli of the shortest ($i = 1$) and longest ($i = 10$) category only confusion with the next longer ($i = 2$) or the next shorter ($i = 9$) perceptual category was accounted for, respectively. With these assumptions we could approximate the processing rates $v(x, i)$ for the 10×10 combinations of stimuli (x) and response categories (i) by only 11 parameters, including the overall processing capacity C . Additional free parameters were two criterion positions k_i (a separate one for the perceptual categories 1 and 10), a parameter for stimulus encoding (μ_S) and a parameter representing residual latency (T_0). Altogether, 15 parameters were used to predict 10 RT distributions for correct identification and the overall probability for correct identification (Figure E-1).

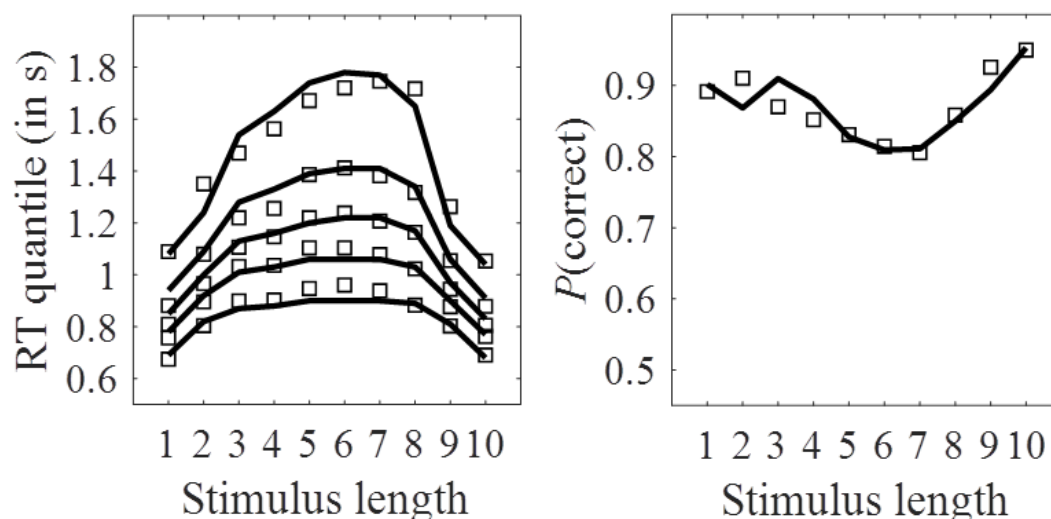


Figure E-1. Data (squares) from Lacouture and Marley (2004, Exp. 2) and model predictions of the Poisson random walk model (unmarked points connected with lines). The left panel are data and model predictions for the 10%, 30%, 50% (median), 70%, and 90% RT quantiles (from bottom to top) in the ten stimulus conditions (1 = shortest stimulus category, 10 = longest stimulus category). The right panel shows observed accuracy of report (squares) and model predictions (unmarked points connected with lines) in the ten stimulus conditions. Parameter estimates are summarized in Table E-1, the resulting processing rates in Table E-2. Adapted by permission from Springer Nature: Springer Perception & Psychophysics, “Choice and response time processes in the identification and categorization of unidimensional stimuli”, by Y. Lacouture & A. A. J. Marley, Copyright 2004.

Despite the restrictions placed on the parameters, the model could explain the shape of all ten distributions of correct identifications sufficiently well and closely approximated the observed accuracy of report. The fit could be further improved by less constraints, the assumption of constant processing capacity is clearly an oversimplification in this case. However, the goal was to obtain a reasonably restricted account of the RT distributions and error rates, thus, the model we fitted may be conceived as a minimal model.

Table E-1.

Parameter estimates for the 10-alternative Poisson random walk model fitted to speeded RT in absolute identification (Lacouture & Marley, 2004; their Exp. 2).

Model parameters													
$\eta(x, C)$	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	C	k_i	μ_S	T_0
56.3	0.92	0.92	0.77	0.73	0.69	0.68	0.68	0.72	0.97	82.8	2.00/ 3.81	6.7	0.297

Note—Perceptual evidence for correct categorization ($\eta(x, C)$, in Hz), bias parameters β_i and processing capacity C (in Hz) determined all processing rates greater than zero. Bias parameter of the last category was a scaling parameter ($\beta_{10} = 1$). Different evidence criteria were allowed for the two outer categories ($k_1 = k_{10} = 2.0$) and the eight inner perceptual categories ($k_{2-9} = 3.8$). Parameter T_0 represents residual latency (in s).

Perceptual evidence for the correct perceptual category ($\eta(x, C) = 56.3$ Hz) and visual processing capacity ($C = 82.8$ Hz) were comparable to Experiment 2, the residual latency ($T_0 = 0.297$ s) was slightly higher and the response of sustained encoding channels lower ($\mu_S = 7$). All these parameters were restricted to be identical across all conditions, thus, the bias parameters accounted almost exclusively for differences between the ten RT distributions. It turned out, though, that it was necessary to include a different evidence criterion ($k_{outer} = 2.0$; $k_{inner} = 3.8$) for the two outermost perceptual categories ($i = 1$ and $i = 10$). Without this additional flexibility, the fit worsens considerably, and parameter estimates become biased and assume implausible values. Starting point variation could also account for this discrepancy, although it does not seem clear why the starting point should be biased in favor of categorizing any given stimulus as the longest or shortest. On the contrary, evidence criteria that are equidistant from the starting point are not generally plausible for non-symmetric experimental designs (i.e., a circular arrangement of report categories as was used in Experiment 2) therefore, we chose the latter specification. On a different note, this application also demonstrates that the restriction of k_i being equal for all categories—a restriction that was plausible in Experiment 2—is for the sake of parsimony and does not constitute a general limitation of the n -choice model.

Table E-2.

Expanded table of processing rates $v(x, i)$ for the Lacouture and Marley data (2004, their Exp. 2).

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}
Stim 1	51.8	31.0								
Stim 2	17.0	51.7	14.2							
Stim 3		22.0	43.2	17.6						
Stim 4			21.9	41.3	19.7					
Stim 5				22.8	38.9	21.1				
Stim 6					22.5	38.2	22.1			
Stim 7						21.6	38.3	22.9		
Stim 8							17.4	40.6	24.8	
Stim 9								11.9	54.4	16.5
Stim 10									26.5	56.3

Note—All processing rates were obtained from a total of 11 free parameters. Empty cells represent inactive Poisson processes (i.e., $v(x, i) = 0$). The sum of each row is the processing capacity ($C = 82.8$).

The bias parameters, which account for variation between conditions, may be conceived as dummy variables that scale perceptual evidence for each categorization in absolute identification. The estimates of the bias parameters revealed a U-shaped pattern suggesting systematic effects of stimulus length on perceptual evidence. In line with this notion, the SAMBA model for absolute identification (Brown, Marley, Donkin & Heathcote, 2008) would explain this U-shaped pattern in bias by accounting for effects such as anchoring and mapping (Brown et al., 2008), providing an explicit account of bias and perceptual evidence (i.e., “a theory of drift rates”). With a modelling account like SAMBA it would be possible to give a more complete picture of performance in absolute identification while also avoiding fitting too many processing rates. The Poisson random walk could be extended to a model of absolute identification, but this is outside the scope of the present article. Rather, this application should demonstrate the general applicability of the random walk model to account for multiple perceptual categories and that its predictions are a close approximation even under very restrictive assumptions.

References

- Ashby, G. F. & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, 21, 93–123.
- Audley, R. J. & Pike, A. R. (1965). Some alternative stochastic models of choice. *British Journal of Mathematical and Statistical Psychology*, 18, 207–225.
- Bacon, W. F. & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, 55, 485–496.
- Banks, W. P. & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception & Psychophysics*, 19, 361–367.
- Blurton, S. P., Kesselmeier, M., & Gondan, M. (2012). Fast and accurate calculations for cumulative first-passage time distributions in Wiener diffusion models. *Journal of Mathematical Psychology*, 56, 470–475.
- Blurton, S. P., Gondan, M., & Greenlee, M. W. (2014). Multisensory processing of redundant information in go/no-go and choice responses. *Attention, Perception, & Psychophysics*, 76, 1212–1233.
- Blurton, S. P., Greenlee, M. W., & Gondan, M. (2015). Cross-modal cueing in audiovisual spatial attention. *Attention, Perception, & Psychophysics*, 77, 2356–2376.
- Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends in Cognitive Sciences*, 11, 118–125.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*, 113, 700–765.
- Brainard, D. H. (1997). The Psychophysics Toolbox, *Spatial Vision*, 10, 433–436.

- Brown, S. & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, 112, 117–128.
- Brown, S. D. & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulators. *Cognitive Psychology*, 57, 153–178.
- Brown, S. D., Marley, A. A. J., Donkin, C., & Heathcote, A. (2008). An integrated model of choices and response times in absolute identification. *Psychological Review*, 115, 396–425.
- Bundesen, C. (1982). Item recognition with automatized performance. *Scandinavian Journal of Psychology*, 23, 173–192.
- Bundesen, C. (1987). Visual attention: Race models for selection from multielement displays. *Psychological Research*, 49, 113–121.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97, 523–547.
- Bundesen, C. (1993). The relationship between independent race models and Luce's choice axiom. *Journal of Mathematical Psychology*, 37, 446–471.
- Bundesen, C., & Habekost, T. (2008). *Principles of visual attention*. Oxford: Oxford University Press.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, 112, 291–328.
- Bundesen, C., & Harms, L. (1999). Single-letter recognition as a function of exposure durations. *Psychological Research / Psychologische Forschung*, 62, 275–279.
- Burbeck, S. L. & Luce, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, 32, 117–133.
- Busmeyer, J. R. & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100, 432–459.

- Busey, T. A. & Townsend, J. T. (2001). Independent sampling vs interim dependencies in whole report processing: Contributions of processing architecture and variable attention. *Journal of Mathematical Psychology*, 45, 283–323.
- Cohen, A. L., & Nosofsky, R. M. (2003). An extension of the exemplar-based random-walk model to separable-dimension stimuli. *Journal of Mathematical Psychology*, 47, 150–165.
- Cox, D. R. (1983). Some remarks on overdispersion. *Biometrika*, 70, 269–274.
- Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes*. London: Methuen & Co Ltd.
- Christensen, J. H., Markussen, B., Bundesen, C., & Kyllingsbæk, S. (2018). A physiologically based nonhomogeneous Poisson counter model of visual identification. *Journal of Experimental Psychology: Human Perception and Performance*, 44, 1383–1398.
- Dai, J., Pleskac, T. J., Pachur, T. (2018). Dynamic cognitive models of intertemporal choice. *Cognitive Psychology*, 104, 29–56.
- Dean, C. B. (1992). Testing for overdispersion in Poisson and binomial regression models. *Journal of the American Statistical Association*, 87, 451–457.
- Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology*, 41, 260–274.
- Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology*, 47, 304–322.
- Ditterich, J. (2006). Stochastic models of decisions about motion detection: Behavior and physiology. *Neural Networks*, 19, 981–1012.
- Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Frontiers in Neuroscience*, 4, 1–24.

- Dubey, S. D. (1970). Compound gamma, beta and F distributions. *Metrika*, 16, 27–31.
- Duncan, J. & Humphreys, G W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Duncan, J. & Humphreys, G W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 578–588.
- Dyrholm, M., Kyllingsbæk, S., Espeseth, T., & Bundesen, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: The case of TVA. *Journal of Mathematical Psychology*, 55, 416–429.
- Dzhafarov, E. N. (1993). Grice-representability of response time distribution families. *Psychometrika*, 58, 281–314.
- Eidels, A., Donkin, C., Brown, S. D., & Heathcote, A. (2010). Converging measures of workload capacity. *Psychonomic Bulletin & Review*, 17, 763–771.
- Eidels, A., Hout, J., W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: facilitory vs. inhibitory interaction in parallel systems. *Journal of Mathematical Psychology*, 55, 176–1908.
- Ennis, D. M. & Bi, J. (1997). The beta-binomial model: accounting for inter-trial variation in replicated difference and preference tests. *Journal of Sensory Studies*, 13, 389–412.
- Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143–149.
- Eriksen, C. W. (1980). The use of a visual mask may seriously confound your experiment. *Perception & Psychophysics*, 28, 89–92.
- Feller, W. (1968). *An introduction to probability theory and its applications* (Vol. 1). New York, NY: Wiley.
- Fisher, D. L. (1984). Central capacity limits in consistent mapping, visual search tasks: Four channels or more? *Cognitive Psychology*, 16, 449–484.

- Habekost, T., Petersen, A., & Vangkilde, S. (2014). Testing attention: comparing the ANT with TVA-based assessment. *Behavior Research Methods*, 46, 81–94.
- Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, 9, 394–401.
- Hockley, W. E. & Murdock, B. B. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review*, 94, 341–358.
- Jones, M. & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review*, 121, 1–32.
- Khodadadi, A. & Townsend, J. T. (2015). On mimicry among sequential sampling models. *Journal of Mathematical Psychology*, 68–69, 37–48.
- Koch, S. H., Weir, C., Westenskow, D., Gondan, M., Agutter, J., Haar, M., Liu, D., Görges, M., & Staggers, N. (2013). Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: a prospective randomized controlled study. *International Journal of Medical Informatics*, 82, 665–675.
- Kvam, P. D. (2019a). A geometric framework for modeling dynamic decisions among arbitrarily many alternatives. *Journal of Mathematical Psychology*, 91, 14–37.
- Kvam, P. D. (2019b). Modeling accuracy, response time, and bias in continuous orientation judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 45, 301–318.
- Kyllingsbæk, S. (2006). Modeling visual attention. *Behavior Research Methods*, 38, 123–133.
- Kyllingsbæk, S., Markussen, B., & Bundesen, C. (2012). Testing a Poisson counter model for visual identification of briefly presented, mutually confusable single stimuli in pure accuracy tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 628–642.

- Lacouture, Y. & Marley, A. A. J. (2004). Choice and response time processes in the identification and categorization of unidimensional stimuli. *Perception & Psychophysics*, 66, 1206–1226.
- Laming, D. R. J. (1968). *Information theory of choice reaction time*. New York: Wiley.
- Leite F. P. & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics*, 72, 246–273.
- Link, S. W. (1975). The relative judgment theory of two choice response time. *Journal of Mathematical Psychology*, 12, 114–135.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–105.
- Little, D. R., Eidels, A., Fific, M., & Wang, T. (2015). Understanding the influence of distractors on workload capacity. *Journal of Mathematical Psychology*, 68–69, 25–36.
- Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, 103, 603–649.
- Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review*, 109, 376–400.
- Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology*, 55, 207–234.
- Logan, G. D. & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, 108, 393–434.
- Luce, R. D. (1963). Detection and recognition. In: R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. 1, pp. 103–189. New York: Wiley.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

- Maddox, W. T. & Ashby, F. G. (1996). Perceptual separability, decisional separability, and the identification-speeded classification relationship. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 795–817.
- McGill, W. J. (1963). Stochastic latency mechanisms. In: R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*. Vol 1, pp. 309–360. New York: Wiley.
- Miller, J. (1982). Coactivation of divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247–279.
- Moran, R., Zehetleitner, M., Müller, H. J., & Usher, M. (2013). Competitive guided search: meeting the challenge of benchmark RT distributions. *Journal of Vision*, 13, 1–31.
- Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47, 90–100.
- Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 87, 691–692.
- Nelder, J. A. & Mead, R. (1965). A simplex method for function optimization. *The Computer Journal*, 7, 308–313.
- Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, 118, 280–315.
- Nosofsky, R. M., & Palmeri, T. J. (1997a). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266–300.
- Nosofsky, R. M. & Palmeri, T. J. (1997b). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception & Psychophysics*, 59, 1027–1048.
- Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 324–354.

- Palmeri, T. J., Love, B. C., & Turner, B. M. (2017). Model-based cognitive neuroscience. *Journal of Mathematical Psychology*, 76, 59–64.
- Petersen, A. & Andersen, T., S. (2012). The effect of exposure duration on visual character identification in single, whole, and partial report. *Journal of Experimental Psychology, Human Perception and Performance*, 38, 498–514.
- Petersen, A., Kyllingsbæk, S., & Bundesen, C. (2012). Measuring and modeling attentional dwell time. *Psychonomic Bulletin & Review*, 19, 1029–1046.
- Pike R. (1973). Response latency models for signal detection. *Psychological Review*, 80, 53–68.
- Pleskac, T. J. & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, 117, 864–901.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160–174.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86, 446–461.
- Ratcliff, R. (1980). A Note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology*, 21, 178–184.
- Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, 88, 552–572.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*, 95, 238–255.

- Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, 9, 278–291.
- Ratcliff, R. (2018). Decision making on spatially continuous scales. *Psychological Review*, 125, 888–935.
- Ratcliff, R., Hasegawa Y. T., Hasegawa, R. P., Smith, P. L., & Segraves, M. A. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *Journal of Neurophysiology*, 97, 1756–1774.
- Ratcliff, R. & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, 20, 873–922.
- Ratcliff, R & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347–356.
- Ratcliff, R. & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16, 323–341.
- Ratcliff, R. & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9, 438–481.
- Schwarz, W. (1989). A new model to explain the redundant signals effect. *Perception & Psychophysics*, 46, 498–500.
- Shepard, R. N. (1957). Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. *Psychometrika*, 22, 325–345.
- Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 591–600.

- Smith, P. L. (1995). Psychologically principled models of visual simple reaction time. *Psychological Review*, 102, 567–593.
- Smith, P. L. (1998). Bloch's Law predictions from diffusion process models of detection. *Australian Journal of Psychology*, 50, 139–147.
- Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, 44, 408–463.
- Smith, P. L. (2016). Diffusion theory of decision making in continuous report. *Psychological Review*, 123, 425–451.
- Smith, P. L. & Corbett, E. A. (2019). Speeded multielement decision-making as diffusion on a hypersphere: theory and application to double target detection. *Psychonomic Bulletin and Review*, 26, 127–162.
- Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, 116, 283–317.
- Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: response time distributions with masked and unmasked displays. *Vision Research*, 44, 1297–1320.
- Smith, P. L., & Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, 53, 293–315.
- Smith, P. L. & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, 32, 135–168.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74 (11, Whole No. 498).
- Stone, M. (1960). Models for choice reaction time. *Psychometrika*, 25, 251–160.
- Tanner, W. P. & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401–409.

- Theeuwes, J. (1990). Perceptual selectivity is task dependent: Evidence from visual search. *Acta Psychologica*, 74, 81–99.
- Townsend, J. T. (1981). Some characteristics of visual whole report behavior. *Acta Psychologica*, 47, 149–173.
- Townsend, J. T. (1992). On the proper scales for reaction time. In H. Geissler, S. Link, & J. T. Townsend (Eds.), *Cognition, information processing and psychophysics: Basic issues*. Hillsdale, NJ: Earlbaum.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, New York, Melbourne: Cambridge University Press.
- Townsend, J. T., Hu, G. G., & Evans, R. J. (1984). Modeling feature perception in brief displays with evidence for positive interdependencies. *Perception & Psychophysics*, 36, 35–49.
- Townsend, J. T. & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, 39, 321–359.
- Townsend, J. T. & Wenger, M. J. (2004). A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003–1035.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Usher, M. & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592.
- van Oeffelen, M. P., & Vos, P. G. (1982). Configurational effects on the enumeration of dots: counting by groups. *Memory & Cognition*, 10, 396–404.
- Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review*, 7, 424–465.

- Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, 7, 208–256.
- Verdonck, S. & Tuerlinckx, F. (2016). Factoring out non-decision time in choice RT data: theory and implications. *Psychological Review*, 123, 208–218.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, 13, 37–58.
- Wenger, M. J. & Townsend, J. T. (2006). On the costs and benefits of faces and words: process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 755–779.
- Wolfe, J. M. (2007). Guided Search 4.0: current progress with a model of visual search. In: (W. D. Gray, Ed.), *Integrated Models of Cognitive Systems*. New York: Oxford University Press.
- Wolfe, J. M., Palmer, E. M., & Horowitz, T. S. (2010). Reaction time distributions constrain models of visual search. *Vision Research*, 50, 1304–1311.
- Zhang, J. & Bogacz, R. (2010). Optimal decision making on the basis of evidence represented in spike trains. *Neural Communication*, 22, 1113–1148.